THE
PROCEEDINGS
OF
THE PHYSICAL SOCIETY

VOL. 53, PART 4 1 July 1941 No. 298

CONTENTS

Price to non-Fellows 7*s*. net; post free 7*s*. 5*d*.
Annual subscription 35*s*. post free, payable in advance

Published by
THE PHYSICAL SOCIETY
1 Lowther Gardens, Exhibition Road
London, S.W. 7

Printed by
TAYLOR AND FRANCIS, LTD.,
RED LION COURT, FLEET STREET, LONDON, E.C. 4

THE PHYSICAL SOCIETY

FOUNDED 1874 . INCORPORATED 1878

OFFICERS OF THE SOCIETY, 1940–41 :

Office of the Society :—1 *Lowther Gardens, Exhibition Road, London, S.W.7.*
Telephone : *Kensington* 0048.

INSTRUCTIONS TO AUTHORS

NOTE. The acceptance of a paper for publication in the Proceedings rests with the Council, advised by its Editing Committee. The high cost of printing renders it imperative to exclude matter that is not novel and not of importance to the understanding of the paper.

Authors offering original contributions for publication in the Proceedings should observe the following directions; failure to comply with these may cause considerable delay in publication.

Manuscript.—A clear and concise style should be adopted, and the utmost brevity consistent with effective presentation of the original subject-matter should be used. The copy should be easily legible, preferably typewritten and double-spaced. It should receive a careful final revision before communication, since alterations are costly when once the type has been set up. Mathematical expressions should be set out clearly, in the simplest possible notation.

References.—In references to published papers the author's name and initials followed by the year and then by the title of the journal in italics, volume and page should be given thus: Vincent, J. H., 1931. *Proc. Phys. Soc.* 43, 157. The abbreviations given in the *World List of Scientific Periodicals* should be employed.

Drawings and tables.—Diagrams must be carefully drawn in Indian ink on white paper or card. Their size and thickness of line must be sufficient to allow of reduction. *Lettering and numbering should be in pencil,* to allow of printing in a uniform style. The number of diagrams should be kept down to the minimum. Photographs of apparatus are not ordinarily accepted. Data should in general be presented in the form of either curves or tables, but not both. Footlines descriptive of figures, and headlines indicative of contents of tables, should be supplied. *Sheets should not be larger than foolscap.*

Abstracts.—Every paper must be accompanied by an abstract in duplicate, brief but sufficient to indicate the scope of the paper and to summarize all novel results.

Proofs.—Proofs of accepted papers will be forwarded to authors. They should be returned promptly with errors corrected, but additions to or other deviations from the original copy should be avoided. When the cost of alterations to proofs, apart from corrections of errors, exceeds 5 per cent of the cost of setting up the type of the paper, the excess must be borne by the author.

Reprints.—Fifty copies of printed papers will be supplied gratis. Extra copies may be purchased at cost price.

Contributions by non-Fellows.—Papers by non-Fellows must be communicated to the Society through a Fellow.

Republication.—Permission to reproduce papers or illustrations contained therein may be granted by the Council on application to the Hon. Secretaries.

INDEX SLIP

OF THE

PROCEEDINGS OF THE PHYSICAL SOCIETY

Vol. 53, 1941, Part 3

SUBJECT INDEX

539.374

de Bruyne, N A
1941.01.06. Note on viscous and plastic flow.
Proc. phys. Soc. Lond. **53**, 251–257 (1941).

546.33 : 535.338.3

Searle, G F C
1941.03.14. The determination of $1/\lambda_2 - 1/\lambda_1$ for the D_1, D_2 lines of sodium.
Proc. phys. Soc. Lond. **53**, 265–271 (1941).

546.621 : 536.413

Wilson, A J C
1940.12.05. The thermal expansion of aluminium from 0° to 650° C.
Proc. phys. Soc. Lond. **53**, 235–244 (1941).

548.73.07

Wilson, A J C
Lipson, H
1940.12.28. The calibration of Debye-Scherrer X-ray powder cameras.
Proc. phys. Soc. Lond. **53**, 245–250 (1941).

621.385.1

Darbyshire, James A
1940.12.05. Initial velocity currents in thermionic valves.
Proc. phys. Soc. Lond. **53**, 219–232 (1941).

621.396.67

Burgess, R E
1940.11.14. Noise in receiving aerial systems.
Proc. phys. Soc. Lond. **53**, 293–304 (1941).

678.1 : 536.21

Schallamach, Adolf
1941.01.15. The heat conductivity of rubber at low temperatures.
Proc. phys. Soc. Lond. **53**, 214–218 (1941).

92 (Atkinson, Ernest Cuthbert)

A[tkinson], H W
1941.05.01. Ernest Cuthbert Atkinson.
Proc. phys. Soc. Lond. **53**, 318–320 (1941).

92 (Hadfield, Sir Rober

Griffiths, Ezer
1941.05.01. Sir Robert Hadfield.
Proc. phys. Soc. Lond. **53**, 316–318 (1941).

92 (Thomson, Sir J J)

Born, Max
Crowther, J A
1941.05.01. Sir J. J. Thomson, O.M., F.R.S.
Proc. phys. Soc. Lond. **53**, 305–316 (1941).

92 (Tomkins, John Arthur)

Andrews, J P
1941.05.01. John Arthur Tomkins.
Proc. phys. Soc. Lond. **53,** 320–321 (1941).

AUTHOR INDEX

92 (Tomkins, John Arthur)

Andrews, J P
1941.05.01. John Arthur Tomkins.
Proc. phys. Soc. Lond. **53,** 320–321 (1941).

92 (Atkinson, Ernest Cuthbert)

A[tkinson], H W
1941.05.01. Ernest Cuthbert Atkinson.
Proc. phys. Soc. Lond. **53,** 318–320 (1941).

537.591

Blackett, P M S
1941.03.05. Cosmic rays : Recent developments. (The twenty-fourth Guthrie Lecture.)
Proc. phys. Soc. Lond. **53,** 203–213 (1941).

92 (Thomson, Sir J J)

Born, Max
Crowther, J A
1941.05.01. Sir J. J. Thomson, O.M., F.R.S.
Proc. phys. Soc. Lond. **53,** 305–316 (1941).

532.66

Brown, R C
1941.02.22. Note on the energy associated with capillary rise.
Proc. phys. Soc. Lond. **53,** 233–234 (1941)

539.374

de Bruyne, N A
1941.01.06. Note on viscous and plastic flow.
Proc. phys. Soc. Lond. **53,** 251–257 (1941).

621.396.67

Burgess, R E
1940.11.14. Noise in receiving aerial systems.
Proc. phys. Soc. Lond. **53,** 293–304 (1941).

92 (Thomson, Sir J J)

Crowther, J A
Born, Max
1941.05.01. Sir J. J. Thomson, O.M., F.R.S.
Proc. phys. Soc. Lond. **53,** 305–316 (1941).

621.385.1

Darbyshire, James A
1940.12.05. Initial velocity currents in thermionic valves.
Proc. phys. Soc. Lond. **53**, 219–232 (1941).

535.668

Ellis, H W
1941.02.24. The technical aspect of colour tolerance.
Proc. phys. Soc. Lond. **53**, 288–292 (1941).

92 (Hadfield, Sir Robert)

Griffiths, Ezer
1941.05.01. Sir Robert Hadfield.
Proc. phys. Soc. Lond. **53**, 316–318 (1941).

535.39

Johnson, B K
1941.01.01. Reflecting-power measurements in the spectral region 2000 A. to 1300 A.
Proc. phys. Soc. Lond. **53**, 258–264 (1941).

548.73.07

Lipson, H
Wilson, A J C
1940.12.28. The calibration of Debye-Scherrer x-ray powder cameras.
Proc. phys. Soc. Lond. **53**, 245–250 (1941).

535.6

Perry, J W
1941.02.15. Colour, its measurement, discrimination and specification.
Proc. phys. Soc. Lond. **53**, 272–287 (1941).

536.21 : 678.1

Schallamach, Adolf.
1941.01.15. The heat conductivity of rubber at low temperatures.
Proc. phys. Soc. Lond. **53**, 214–218 (1941).

535.338.3 : 546.33

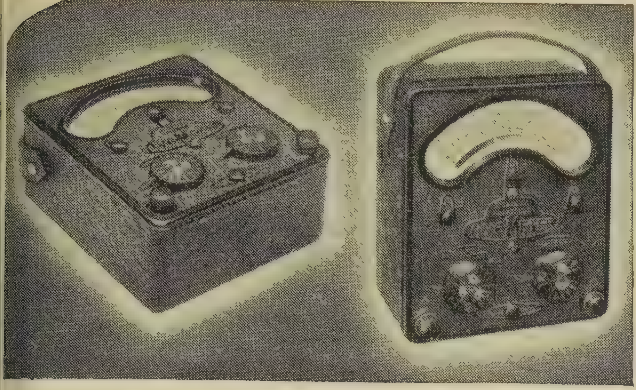
Searle, G F C
1941.03.14. The determination of $1/\lambda_2 - 1/\lambda_1$ for the D_1, D_2 lines of sodium.
Proc. phys. Soc. Lond. **53**, 265–271 (1941).

536.413 : 546.621

Wilson, A J C
1940.12.05. The thermal expansion of aluminium from 0° to 650° C.
Proc. phys. Soc. Lond. **53**, 235–244 (1941).

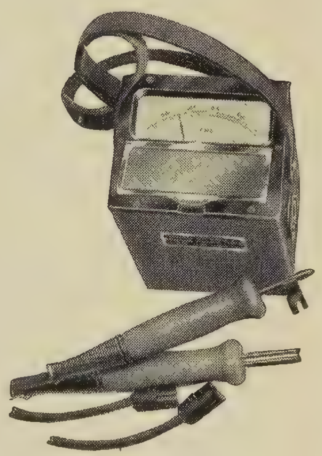
548.73.07

Wilson, A J C
Lipson, H
1940.12.28. The calibration of Debye-Scherrer x-ray powder cameras.
Proc. phys. Soc. Lond. 53, 245–250 (1941).



46-range Model 7 Universal AvoMeter

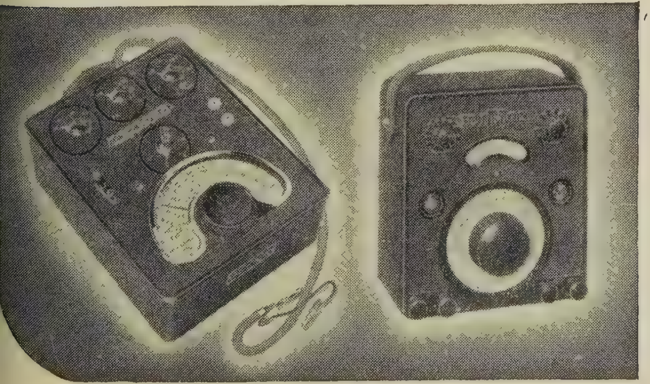
40-range Model 40 Universal AvoMeter

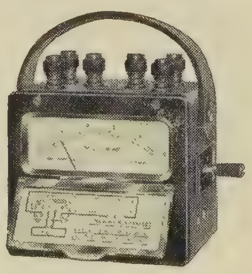


" AVO " Bonding Meter

All-wave " AVO " Oscillator

" AVO " Test Bridge





" AVO " Milliohmmeter


Taylor & Francis
LIMITED
SCIENTIFIC & COMMERCIAL
LETTERPRESS PRINTERS
PUBLISHERS AND
BOOKBINDERS
TELEPHONE: CENTRAL 5191
PRINTERS FOR
OVER A CENTURY
RED LION COURT
CRANE COURT
FLEET ST., LONDON, E.C.4

THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 53, PART 4 | 1 July 1941 | No. 298

THE SENSITIVE FLAME

BY E. N. DA C. ANDRADE, PH.D., D.SC., F.R.S.,
Quain Professor of Physics in the University of London

Twenty-fifth Guthrie Lecture, delivered 4 April 1941

§ 1. INTRODUCTION

I AM deeply sensible of the honour which the Society, of which I am proud to be a Fellow, has done me by asking me to deliver this lecture, to which the honoured name of Guthrie is attached. It was with some diffidence that I accepted the invitation, for to follow those who have given the lecture in previous years might well give pause to men of mightier metal than myself. Lean fare, however, is the rule in war, and I feel that, whereas in normal times you might expect an intellectual feast, your demands in these days will be more modest, and you may perhaps be disposed to be satisfied with the simple dish that I can offer. My choice of a subject has been, in any case, considerably restricted by the total destruction of my laboratory with most of my papers and all my slides.* I am, therefore, bold to come before you with a simple theme, and I know that from an audience where I see so many old friends I can hope for indulgence.

Guthrie was interested in the subject of sound and, in fact, was the first to draw attention to the attraction exercised by a vibrating prong on light bodies, an effect to which much attention was afterwards given and which I myself had to study in some work which I did on Kundt's tube. Now in 1858, when Guthrie was about twenty-five years old, a certain Professor Leconte noticed a gas flame at a concert which exhibited pulsations which were synchronous with what he called "audible beats"—certain features of the music, let us say. This was the phenomenon afterwards so well known as the sensitive flame, a phenomenon which Guthrie must often have showed and which certainly nearly everyone here must have demonstrated at one time or another. Owing to its beauty and the simplicity of the apparatus needed, it is shown every year to a host of students, but, strangely enough, although the phenomenon is nearly eighty years old, the explanations which accompany it are either indefinite or demonstrably wrong.

* The slides shown during the lecture I owed to the good offices of the Honorary Secretary for Business, who was instrumental in having them prepared from surviving prints.

It was in the hope of finding an explanation of what is a puzzling phenomenon that I started investigations which continued at intervals over some years. The results of these investigations I am about to lay before you, relying, to retain your interest, on the fact that the flame may have teased your curiosity at some time or other. The investigations have led me in one or two places into rather wider fields than would at first have seemed relevant. I must not weary you with further excuses, but begin my story.

§ 2. PREVIOUS WORK ON THE SENSITIVE FLAME

Our knowledge of the sensitive flame up to the end of the last century is exposed in the second volume of Lord Rayleigh's classic work, but it may be well, perhaps, to summarize very briefly the position from which I started. The name " sensitive flame " is only adventitiously attached to the phenomenon, which can be demonstrated in a less complicated aspect with jets of air, distinguished by admixture of smoke, in air; or jets of coloured water in water. The problem is really that of the temporary instability, provoked by regular vibrations in the medium, of steady jets of fluid projected into the same fluid. Surface tension, originally invoked as an explanation, plays no part at all. As regards the fundamental experimental facts, Tyndall considered that he had shown that the sound is only effective if it falls on the root of the flame. This is, however, not strictly correct, but that the sound is much more effective at this spot than it is if applied higher up the jet is true and significant.

Another fundamental fact was demonstrated by Rayleigh, namely, that it is variations of amplitude, and not variations of pressure, that produce the instability : the flame is unaffected at the nodes, and becomes turbulent at the loops, of a stationary wave. Rayleigh, who introduced the method of stroboscopic observation, gave it as his opinion—" the answer is still, perhaps, in some cases open to doubt "—that the phenomenon was one of increasing sinuosity, not increasing varicosity,* a view which the present work bears out. He also stated that the kinematic viscosity of the fluid was the factor that determined the maximum velocity for non-turbulence and the region of frequency within which the jet is sensitive, as can be demonstrated by consideration of dynamical similarity.

Recently Dr. G. B. Brown, working in my laboratory, has used the stroboscopic method to investigate the periodic motion of smoke-laden air jets into air and has confirmed and extended Rayleigh's observations (1884) that the state of instability produced by a sound is not irregular turbulence, but strictly periodic motion. The sinuous motion is accompanied by the development of vortices to either side. The turbulence produced by increasing the velocity of the jet above what is compatible with laminar motion is, in contradiction to what is often stated, quite different from the effect produced by the sound.

The flame is generally said to be sensitive only if on the point of flaring, but

* The name given by Rayleigh to the form of a figure of revolution exhibiting alternate swellings and contractions.

a more precise statement is that the sensitiveness increases rapidly with the nozzle velocity of the jet, and the highest velocity that is used is that just below the point of turbulence.

§ 3. EXPERIMENTS WITH THE SENSITIVE FLAME

Considerations to be discussed later have convinced me that the phenomenon is due simply to a periodic movement, normal to the axis of the jet, of the air, relative to the nozzle from which the jet issues; a movement which can be brought about either by a sound wave travelling normal to the jet axis or by vibration of the nozzle normally to its axis. If this is so, a sound wave travelling in the direction of the axis of the jet should have no effect. The establishment of this is another experiment fundamental for the understanding of the phenomenon.

A sensitive flame from a circular steatite nozzle of about 1 mm. diameter was set up on a tall slender support, and a loudspeaker, driven by a valve set and amplifier which generated pure tones of a wide range of possible frequencies and intensities, was placed horizontally opposite to the nozzle, so as to send waves normally to the jet axis. The range of frequencies and intensities to which the flame was very sensitive was noted. The loudspeaker was then arranged above the flame, at the same distance from the nozzle as before, so as to send sound waves along the axis of the flame. Frequencies and intensities in the range which previously produced marked flaring of the familiar type now had no effect. The variation of frequency showed that this behaviour was not due to the orifice being at the node of a stationary wave formed by reflection at the floor : varying the distance of the loudspeaker confirmed this. The experiment was carried out in a large bare room, to avoid considerable reflections from walls or furniture.

The Tyndall sensitive flame is best obtained by supplying coal gas, unmixed with air, to a burner fitted with a steatite nozzle pierced by a circular hole about 1 mm. in diameter. The tall flame consists of an extended cone of unburnt gases, which is from 20 to 35 cm. high when the flame is in the sensitive condition, surrounded by a zone (or zones) of burning gas which is luminous in its upper part, but not close to the nozzle. I find that if the diameter of the nozzle is large, say 8 mm., no well-defined inner cone is formed and the flame is not sensitive : a sufficiently rapid rate of supply of gas leads to the formation of an inner cone, but produces flaring. If the diameter of the nozzle is very small, any attempt to produce a long flame leads to flaring.

As regards the reason for the inner cone, it is observed that when it is formed the burning gas is separated from the nozzle by a distance of the order of the radius of the orifice. For the case of a circular nozzle with an infinite baffle the stream lines of the air in the neighbourhood of the flame are parallel to the plate, and so normal to the jet, in the neighbourhood of the orifice, as established mathematically by Schlichting (1923) and experimentally by Andrade and Tsien (1937) (see Andrade and Tsien, figure 9, for a detailed drawing). This represents the case when the inertia effects can be neglected. If the velocity of efflux of the flame

gases is high, the velocity of the air will be high, and owing to its inertia it will be carried into the flame at the base of the flame and will give rise to the cone of unburnt gas, as in a Bunsen flame. As the critical velocity for flaring is, to a first approximation, inversely as the diameter, it is clear that diminishing the diameter, while maintaining the flame in the sensitive state by increasing the velocity, will lead to the formation of an inner cone of unburnt air-gas mixture.

This cone is the sensitive part of the flame, the luminous mantle being merely an indicator. This can be shown by introducing a very fine platinum wire into the flame. As long as it does not touch the inner cone, whether it be in the upper or the lower part of the flame, no disturbance is produced. If it touches the inner cone the flame flares, the disturbance becoming more violent as the wire approaches the nozzle. The disturbance produced by introducing the wire into the inner cone near the base of the flame resembles superficially that produced by sound. The wire, be it noted, produces turbulence at the base of the flame.

In confirmation of the view that it is the cone which is the sensitive jet we may cite not only that the coal-gas flame is not sensitive when there is no cone, but also the fact that the hydrogen flame, which never has a cone, is not sensitive. By adding coal gas to hydrogen, however, an inner cone can be formed and a sensitive flame produced.

Another fact which, as will be seen later, is of considerable importance is that, as I have shown with orifices of various dimensions and proportions, a flame from a uniform slit orifice, a flat flame, is not sensitive to anything like the same extent as a flame from circular orifice.

So much for the flame: we will now turn to less complicated jets.

§4. THE WATER-INTO-WATER JET: CIRCULAR ORIFICE

Since temperature gradients, inhomogeneities of density and suchlike make the case of the flame very complicated, the experiments designed to investigate the cause of sensitiveness were carried out with jets of water into water.

The disposition of the apparatus is shown in figure 1. The jet proceeded horizontally from a nozzle of the form shown at (*a*), carefully drawn from glass tube so as to have a circular orifice. The water through which it flowed was contained in a trough 3 cm. deep, with a glass bottom of area 60×35 cm. The edges of the trough were beached to damp out ripples. Under it was a mirror at 45°, for purposes of illumination. The water to the jet, coloured with Kiton red, came from a reservoir carried by a wooden framework attached to that supporting the mirror box. The supply tube was furnished with a water-calibrated rotameter, to measure the flow. The mirror box stood on a massive slab of slate, which rested on loose cotton swabs carried by a second massive slate slab separated from the floor by cotton swabs. The trough rested on slabs of Sorbo rubber carried by the mirror box. These precautions were necessary in order to shield the trough as much as possible from vibration.

The methods of illumination were various:

(1) An approximately parallel beam from a 60-ampere arc, with a diffusing screen under the trough, for short-exposure photographs showing the jet as it appeared to the eye. The arc was enclosed in a brick-walled space (an empty fireplace) fronted with plate glass, to avoid disturbances by the sound.

(2) Stroboscopic illumination, the light from an arc being interrupted by a rotating wheel furnished with slits, with the usual arrangement of lenses. Arc and wheel were enclosed, as above.

(3) Spark photography, with a spark between magnesium spheres amalgamated with mercury, from jars of capacity about 1μF., charged by an electrostatic machine.

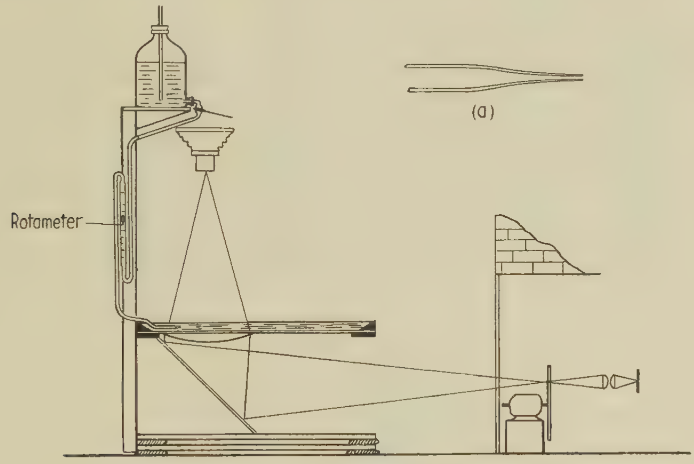


Figure 1.

The jet was photographed with an 8-inch Dallmeyer f 2·9 lens, with Ilford panchromatic plates of from 1000 to 8000 H & D.

For low disturbing frequencies a motor with a slightly unbalanced flywheel was used, which covered a range from the lowest frequency up to about 25 ∼. For higher frequencies a loudspeaker driven by a set giving a pure note was used: with this, frequencies as high as 10 000 and more, and as low as 15 ∼ could be satisfactorily produced.

The undisturbed jet is very fine and sharp. It is particularly sensitive to sound: a picture of the large symmetrical disturbance produced by a comparatively faint sound was published some time ago (Andrade, 1936).

The response of the jet depends upon three factors: the velocity of efflux of the fluid, the frequency of the disturbing sound and the intensity (amplitude) of the sound. The general effect of the frequency is masked by spurious resonance effects, discussed later, for I have succeeded in showing that the sharp response of the jet at certain selected frequencies is due to resonant response, either of structural members of the scaffolding supporting the tube from which the fluid

issues, or of the air enclosed in the room. The very large increase of response as a particular frequency, within a very narrow range of frequencies, is approached may be taken, then, as an effect of increased amplitude of disturbance. Plate 1 shows a jet, from a circular orifice of diameter 0·325 mm., disturbed by vibrations of increasing amplitude, obtained by altering the frequency of a note of approximately constant intensity in steps from 321 $\sim$ to 325 $\sim$, the last being a resonant frequency. The photographs were taken with an exposure of about 1/16 second, and show the jet as it appears to the eye. At small amplitude the jet shows a slight thickening remote from the orifice. As the amplitude increases this thickening retreats towards the orifice and manifests itself as a separation into two, or forking, of the jet. At medium amplitudes the jet appears to reunite, but at larger amplitudes the arms of the fork diverge more and more. The forking is followed by turbulence, as seen in the last four photographs.

The effect of increasing the rate of efflux of the jet is similar to that of increasing the amplitude of the disturbance, to illustrate which we may consider some pictures, taken with stroboscopic illumination, which bring out other points as well. In plate 2, *a* and *b* represent a jet disturbed by a sound of constant intensity and frequency 124 $\sim$, issuing from the same orifice at rates of 295 cm^3/h. and 300 cm^3/h. respectively (Reynolds number $R = 161$ and 164 respectively *). They show first of all how the effect of the vibration is to cause a sinusoidal disturbance which increases in amplitude until it disengages alternate vortices to either side, forming a street which differs from a Kármán street in that, firstly, the vortices are very short in the direction normal to the plane of the paper and, secondly, they do not have the right spacing. It appears that when b/a ($a =$ distance between successive vortices measured parallel to jet, $b =$ spacing measured normal to jet) is small, as it is initially, the two sides of the street diverge, but for a certain value of b/a there is a stable parallel vortex street, as with extended linear vortices. This would account for the phenomenon shown in plate 2, *a*, since with the slowly increasing separation, the street never being very far from parallel, the stable system has an opportunity of establishing itself. When there is a large disturbance, however, and rapidly increasing separation, there is never any approximation to parallelism, and the two separate series of vortices travel independently of each other, as shown in plate 2, *b*.

The early stages of the development of the disturbance are best shown by using stroboscopic illumination with a " sound " of very low frequency, 9·8 $\sim$ in the case of the jet shown in plate 3, *a*, *b* and *c*, which correspond to rates of efflux of 315, 340 and 350 cm^3/h. respectively ($R = 172$, 186 and 191), with a fixed intensity of sound, and also represent the kind of effect we get by increasing the amplitude of the disturbance at constant rate of efflux. The shedding of vorticity from the extremes of the disturbed jet gives a beautiful pattern when the disturbance is large.

With sound of greater amplitude, forking can be obtained at this low frequency,

* The linear dimension used is the radius of the orifice. See Andrade and Tsien, loc. cit.

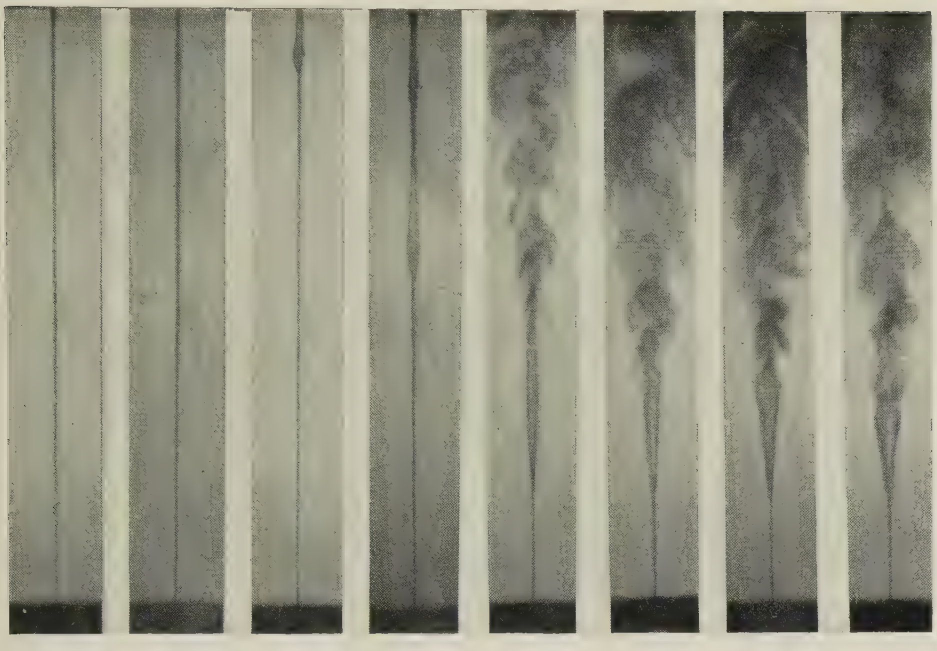
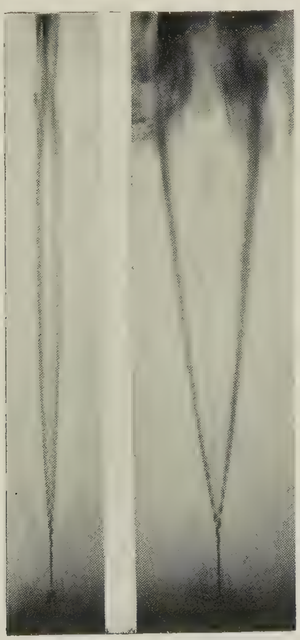


Plate 1. Water jet : increased response caused by increasing amplitude of disturbance.



a *b*

Plate 2. Forking of water jet.

To face page 334

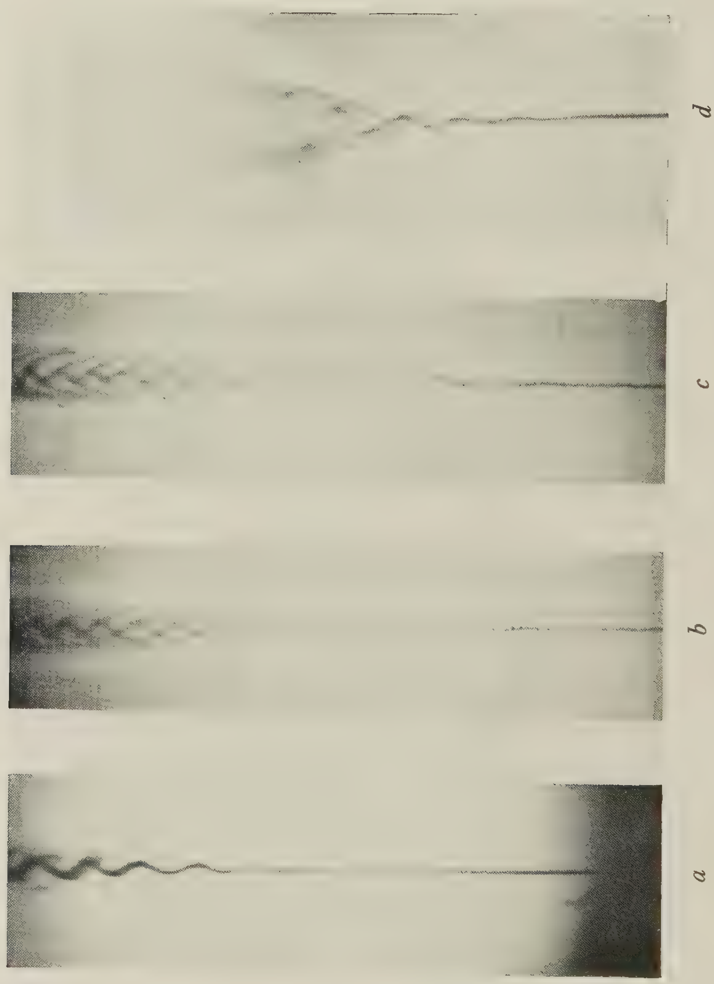


Plate 3. Water jet : early stages of disturbance with low frequency and stroboscopic illumination.

as shown in plate 3, *d*, which represents an outflow of 345 cm^3/h. ($R = 188$). This phenomenon of forking is well known with the sensitive flame, and is illustrated, for instance, in figure 5, reproduced to illustrate another point.

§ 5. THE QUESTION OF RESONANCES

One of the most remarkable phenomena shown by the circular water-jet is the sharp response to particular frequencies. This is illustrated by the photographs of plate 1, the circumstances of which have been already described. S. R. Humby (1927) and G. B. Brown (1932) described a similar selective response with the sensitive flame. Humby was unable to find any relationship between the various frequencies of selective response. Although Brown found that the particular frequencies at which the flame responded were the same for a wide range of different orifices, nevertheless he decided that the selective response "expresses a property of the gaseous jet and not one due to causes inherent in the apparatus used", and stated that these frequencies are constants for any particular gas. He considered the possibility of reflexions from different parts of the room (stationary waves) as a cause of the selective response, but decided that he had eliminated it by finding the same frequencies in another room. He gave a theoretical explanation of the response.

Nevertheless, it is hard to see how a jet can possibly possess of itself, as he contended, a series of resonant frequencies, and experiments were therefore carried out to find the cause of the sharpness of response. In the first place a determination was made, over a very wide range of frequencies, of the particular ones which produced marked response. Periodic disturbances of low frequency were produced by the motor carrying a heavy unbalanced flywheel, which was placed on the floor near the apparatus. The oscillator and loudspeaker carried the range continuously on to somewhere near the limit of audibility. A series of more or less equally spaced resonant frequencies was found, some of which were much more effective than others: the following list of 54 frequencies is fairly complete within the range covered.

Table 1. Resonant frequencies (cycles/sec.) with a jet 0·325 mm. in diameter

1·405	5·8	12·3	160·0	615·0
1·82	6·3	13·6	207·0	631·0
2·85	6·8	15·1	227·0	700·0
3·45	7·5	16·0	258·5	760·0
3·65	8·0	25·0	310·0	825·0
3·8	8·4	43·0	340·0	890·0
4·0	9·0	70·0	400·0	958·0
4·31	9·6	89·0	450·0	1025·0
4·64	10·3	94·0	478·0	1110·0
5·2	11·0	110·0	515·0	1165·0
5·48	11·6	135·0	590·0	

These frequencies for maximum response were not altered by diminishing the rate of flow of the jet, which merely made the jet less sensitive and the response

less marked. They were not altered by varying the depth of liquid in the trough, or the width of the trough. Changing the viscosity of the medium did not modify the frequencies, as was shown, firstly, by altering the temperature of both the jet and the surrounding water from 4° to 40° c., which gives a range of viscosity of 2·4 to 1, and secondly, by using hexane, coloured with iodine, into hexane, which has a kinematic viscosity less than half that of water. The size of orifice also had no effect. Not only were the frequencies not shifted, but they all appeared under the variety of circumstances described.

This was held to prove that the resonant frequencies must be either those of the room or those of certain parts of the apparatus, and attempts, based on this supposition, were made to modify the frequencies. The wooden framework, supporting the reservoir, supply tubes and nozzle, was loaded at certain points with a 30-lb. weight. In the range 0 to 200 cycles this very much diminished the response, as evidenced by the fact that a higher rate of flow was required to make the response evident. Certain of the frequencies were completely suppressed and new frequencies were introduced, but some remained unchanged. Clamping the tube just short of the orifice more firmly to the heavy framework diminished the sensitiveness and shifted the frequencies in the region of a few cycles/sec. These results are consistent with the view that many of the resonances are due to the modes of vibration of the supporting structure and of the free tube which terminates in the orifice, while others, mostly in the range from 50 cycles upwards, are due to the floor and walls of the room and the contained air, the resonances of which lie in about this range (Constable, 1936). These latter would not be shifted by loading the structure, but their effectiveness would be diminished.

Everything in the present investigation is consistent with the supposition that the whole of the effects of sound or other vibrations are due to a transverse relative motion of orifice and surrounding fluid. The effects should therefore be produced by deliberately vibrating the orifice, the medium being stationary. With this in view the approach tube was firmly connected by a short rod to the diaphragm of an Amplion loudspeaker unit, so arranged as to move the orifice of the horizontal jet in a vertical line. If the jet itself had any selective response it should manifest itself plainly as the range of frequencies was traversed. Actually, just the same type of phenomena were produced in this case as in that of disturbance by external sound, but there was no sign of any selective frequency effect. Some control of the amplitude of the sound was necessary to demonstrate the phenomena equally clearly over the whole range of frequencies, but there was nothing in the way of special response at any particular frequency.

A similar experiment was tried with the sensitive flame. By vibrating the jet transversely it was shown that the response varied continuously with frequency, without sign of any particularly favoured frequencies, even down to low notes. Since the amplitude was not measured, the nature of the gradual variation of response has no particular interest.

Another experiment with the horizontal jet served to demonstrate both the lack of selective response and the fact that the orifice is the seat of the sensitiveness. A small disc, 4 mm. in diameter, was fixed by a rod to the diaphragm of the speaker unit so that it could be set in vibration at any frequency desired. It was arranged just beneath the surface of the water, so that it could produce waves propagated in a vertical direction. When just above the orifice, it produced the usual type of disturbance, and variation of the frequency caused no selective response, but merely a gradual change of magnitude which could be counterbalanced by continuous control of the amplitude. When the disc was arranged either at any place above the approach tube, or above any part of the jet except that in the neighbourhood of the orifice, no effect was produced.

Another striking phenomenon is easily explained on the view here put forward. When the disturbing sound is produced either by the unbalanced motor or loud-speaker, the plane in which the broadening of the jet takes place changes at certain frequencies from the usual (horizontal) plane to a plane normal to it. The change of plane, horizontal to vertical, and back again, takes place within a small range of resonant frequencies. If the particular frequencies of selective response are due to vibrations of the nozzle tube, and this is not quite symmetrical about its axis, the result is at once explained.

The results of this investigation of selective response may seem obvious, but the phenomenon is very striking, and since it has misled previous workers it seems worth while to have established its nature, more especially as Rayleigh's theory of the instability of an idealized jet, which we have to consider, suggests that the instability is greater for a particular wave-length.

The sensitive water-jet or sensitive flame has, then, no selective response in itself. It is, however, when suitably disposed, a very delicate detector of vibrations of any structure with which it may be in sound-bearing connection, and it may possibly have applications in this connection. For instance, for investigating the frequencies of a room, a sensitive jet with a heavily weighted nozzle and heavy framework would be a convenient indicator.

For investigation of the causes of the phenomenon it is clear that it is best to vibrate the nozzle and not the medium.

§ 6. LORD RAYLEIGH'S THEORY: THE SINGLE VORTEX SHEET*

Many of the phenomena which have to be explained have now been laid before you. We now turn to the question of their explanation, and will first consider briefly Lord Rayleigh's calculations.

In his theoretical work on the sensitive flame, Rayleigh (1896) considers the two-dimensional motion of an idealized jet, idealized not only in the sense that the liquid is perfect, but also in the sense that the distribution of velocity does not attempt to approximate closely to that in any real jet. The case worked out in most detail is that of a jet consisting of two adjacent sheets in laminar motion,

* Lord Rayleigh, *Theory of Sound*, vol. II, chapter XXI, and *Collected Papers*.

in each of which constant vorticity, of equal magnitude but opposite sign, prevails. The velocity thus has a fixed value V in the x direction along the plane of contact of the two laminae, diminishing steadily to zero value at a distance b to either side of the central plane, and being zero everywhere outside the laminae. The distribution of velocity is shown in figure 2 (a) and the distribution of vorticity in figure 2 (b). Rayleigh proves that, if the jet is subjected to a small periodic disturbance $f(y)e^{i(nt+kx)}$, the disturbance grows exponentially if kb is below a certain value, i.e. we have instability for disturbances of wave-lengths above a certain value, and he further shows that the instability is greater for a certain wave-length. Since the fluid is assumed to be ideal, the results do not necessarily apply to real fluids, where viscosity exercises a stabilizing influence: Rayleigh himself points out " there is therefore ample foundation for the suspicion that

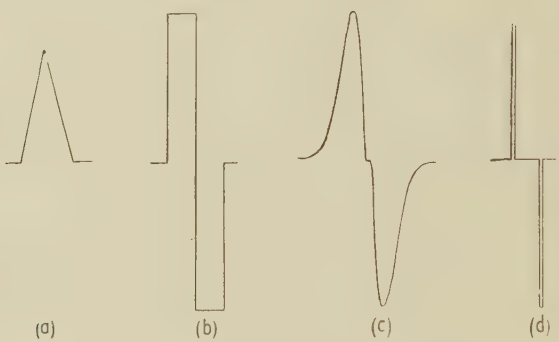


Figure 2.

the phenomena of sensitive jets may be greatly influenced by fluid friction, and deviate materially from the results of calculations based upon supposition of discontinuous changes of velocity ".

Quite apart from any lack of correspondence with reality which may be involved in the arbitrary distribution of velocity and the lack of viscosity, the calculations do not consider certain fundamental features of the sensitive jet from a circular aperture, namely, that the seat of the sensitiveness is the orifice of the jet and that an extremely small amplitude of the motion is required to produce the instability. Again, I have found experimentally that the jet from a slit, which corresponds closely to the two-dimensional case considered by Rayleigh, does not show any such sensitiveness as does the circular jet.

A further experiment that seems to prove that the mechanism contemplated by Rayleigh is not sufficient to explain the sensitive flame concerns the case of a single vortex sheet, produced by letting a wide stream of water, with constant velocity, flow past a wide expanse of stationary water. This gives a good approximation to the case, treated by Rayleigh: " on the upper side of a layer of thickness b the undisturbed velocity U is equal to $+V$, and on the lower side to $-V$, while inside the layer it changes uniformly." * The type of instability which he finds

* *Sound*, vol. II, p. 392.

here is exactly the same as that for the flat jet ; for all wave-lengths, if b is infinitely small, and for a range of wave-lengths if b is finite, the surface is unstable on the criterion adopted throughout the investigation. Experimenting with a single vortex sheet I have, however, found that while the surface becomes unstable, with vortex formation, if the velocity exceeds a certain value, there is no sensitiveness at all to sounds of any frequency within the audible range, or lower.

The vortex sheet was produced by letting a wide stream of water run through a trough, one side of the stream being against the wall of the trough, and so stabilized, while the other side flowed past the stationary water. The trough was 3 cm. deep, 90 cm. long and 26 cm. wide. At one side water from a tank containing a heater * was admitted through an opening 3 cm. wide: it passed through a series of parallel guides to ensure uniform flow. At the far end of the trough was a series of exit pipes with adjustable outflow: all except those opposite the 3-cm. strip were closed during an actual flow experiment. The water in the trough was coloured with 1 in 10 000 Kiton red. It was found that by suitably adjusting the outflow a uniform parallel strip of moving clear water could be obtained, with a sharp boundary, which constituted the approximately two-dimensional single vortex sheet.

Over a wide range of velocities the boundary showed a sinuosity which at high velocities of flow tended to "feather" and form vortices. † Rosenhead (1931) has carried out calculations on the form of the instability of such a vortex sheet, and the forms of the observed instabilities were much as depicted in his diagrams, so that the experiment probably realized closely the desired distribution of velocity. The effect of sounds of various intensities, in some cases considerable, and frequencies was tried, both by having a loudspeaker in various positions, close to the origin of the sheet and elsewhere, and by attaching the unbalanced motor to the framework. No kind of response in any way resembling that of the sensitive flame could be obtained.

The conclusion is, then, that while Rayleigh's investigations indicate that single and double sheets of vorticity are unstable for certain ranges of frequency, his theory gives no indication of the sensitiveness to small periodic disturbances peculiar to the circular jet, and, therefore, has no real application to the sensitive flame. Any ordinary small disturbances, such as are inevitably taking place all the time, are equivalent to trains of all wave-lengths, so that the investigation seems really to point to the inherent instability of all such sheets. The method applied, in fact, of finding whether a small applied deformation grows exponentially or no is that normally used for investigating inherent instability. Further, the single vortex sheet is a good example of the cases treated by Lord Rayleigh, and it is not sensitive in the smallest

* Inserted in order to be able to ensure that the temperature of the water from the tank was the same as that in the trough.

† The apparatus was set up for photography of the various stages of the instability, produced by varying the velocity, when the war broke out, but the photographs were never taken and the apparatus has since been destroyed by enemy action.

degree. In practice ordinary instability occurs only when the strength of the vortex sheet exceeds a certain value, a fact which needs the invocation of viscosity for its explanation. The stabilizing influence of viscosity in many cases is well known : reference need only be made to G. I. Taylor's investigations on the behaviour of a fluid between rotating concentric cylinders.

The question which we have to consider, however, is not one of inherent instability but of instability due to a particular form of disturbance, which only continues so long as that disturbance is acting.

§ 7. THE GROWTH OF THE INSTABILITY OF A PLANE JET

Having come to the conclusion, from a study of the forking to which I have alluded and from other considerations, with which I need not trouble you, that the peculiar form of break-up of the sensitive jet is merely due to the action of two opposite sides of the slender conical vortex sheet, to which it is roughly equivalent, on one another, I was confronted by the question as to how to confirm this belief. It is clear that the three-dimensional case is extremely difficult to handle, and that it is more practicable to study a two-dimensional case—two parallel vortex sheets of opposite vorticity, which can be approximately realized by a jet from a slit orifice. Although, as I have said, such a jet is nothing like as sensitive as a jet from a circular orifice, typical break-up can be realized by making the orifice vibrate transversely to its length.

The point to be investigated is the form assumed by the jet rendered unstable by a periodic transverse relative motion of nozzle and medium. Rayleigh's investigation tells us nothing about the way in which the instability grows. Rosenhead (1931), in his investigation on the single vortex sheet, assumes a sheet consisting of equally spaced parallel line vortices, and then calculates, step by step (11 steps for 8 vortices per wave-length, 4 steps for 12 vortices per wave-length), the displacement of the system, by finding the velocity of each vortex produced by the other vortices of the system, and so the new position, after a given short time interval. He shows that a sinusoidal vortex surface tends to roll up at each crest like a breaker, and gives diagrams depicting the successive stages, to which we have already referred.

The single vortex street is not suitable for detailed comparison with experiment, and is, in any case, too remote from the jet to furnish much help towards an explanation of the latter. Further, it is difficult to apply a suitable periodic disturbance to a single sheet. It was therefore decided to work out the case of the double vortex street, which is equivalent to a two-dimensional jet. The motion has to be followed step by step; instead of doing this by detailed calculations, such as were carried out by Rosenhead, a method based on the hydro-magnetic analogy was adopted.

The distribution of magnetic force round a line current I is similar to the distribution of velocity round a line vortex of strength Γ, the lines of constant force and velocity being circles, and the magnitudes of force and velocity being

$2I/r$ and $\frac{1}{2\pi}\Gamma/r$ respectively. The distribution of magnetic force near any system of line currents must therefore be similar to the distribution of velocity round a similar system of line vortices, and the strength of the magnetic field at any point must be proportional to the velocity at a similar point. Hence if at the position of any one current we measure the magnetic field due to the other currents it will give us the velocity with which the corresponding vortex is being displaced by the other vortices. Having found the field at every line current, we are able to displace the currents to new positions corresponding to the position which the vortices would take up, under their own field, after a small fixed time interval. With this new disposition of the vortex system the field at each current can then be again determined, and a new disposition of the system found. In this way the whole history of the vortex system can be built up in steps. It must be emphasized that a similar step-by-step procedure is all that is possible if purely mathematical methods are adopted. Viscosity and inertia effects are, of course, neglected, as usual in such investigations. In experiments with comparatively slow rates of flow and low frequencies the inertia effects are not likely to be large.

The experimental determination was carried out in the following way. The linear currents were arranged initially in two parallel sine curves, to represent the disturbed position of the parallel jet. For the mathematician who may later carry out calculations on the double vortex sheet the interesting case is that of two infinite sheets: on the other hand, the jet corresponds, as a first rough approximation, to a double sheet infinite in one direction * only. A compromise between the two cases was made by taking three wave-lengths, and considering the behaviour of the mid-one. Actually, if, as in the experimental case, the disturbance is at the orifice, the portion of the jet in which the disturbance is proceeding travels out and is subjected, as it travels, to the effect of a more disturbed portion in advance and a less disturbed portion coming in from behind, so that the disposition adopted gives a kind of average value, adapted to show the general way in which the disturbance develops. This is all that was sought.

The conductors consisted of straight copper rods, 8 mm. in diameter and 186 cm. long, arranged initially in two parallel sine curves, 7 cm. apart, of wave-length 60 cm. and amplitude 3 cm., there being 20 wires per wave length, a much closer distribution than that adopted by Rosenhead. The rods were supported at the top and bottom by horizontal wooden boards pierced with holes, in which they just fitted, but, as it is difficult to drill wood with the accuracy of position required, two subsidiary plates, of stout seasoned celluloid, were bored with great care and used to space the rods at the middle, where measurements were made. These plates can be seen in the photograph of figure 3, which shows the complete installation. The rods were connected top and bottom by short lengths

* Actually the plane jet corresponds more closely to two diverging sheets of which the vorticity falls off directly as the distance from a point within the nozzle, as will be shown later.

of stout cable in such a way that the current flowed up in all the rods on one side and down in all those on the other, as in a flattened solenoid. They are represented by dots in figure 4, which shows the experimental disposition.

To measure the magnetic field a search coil S, 7 cm. high and 8 mm. wide, was constructed of 392 turns of 44 s.w.g. copper wire. It was mounted on a narrow T-shaped stand, the base of which carried a plug, accurately centred with the axis of the coil, of the same diameter as the copper rods. It was provided with two mirrors, at right angles to one another, one being accurately parallel to the

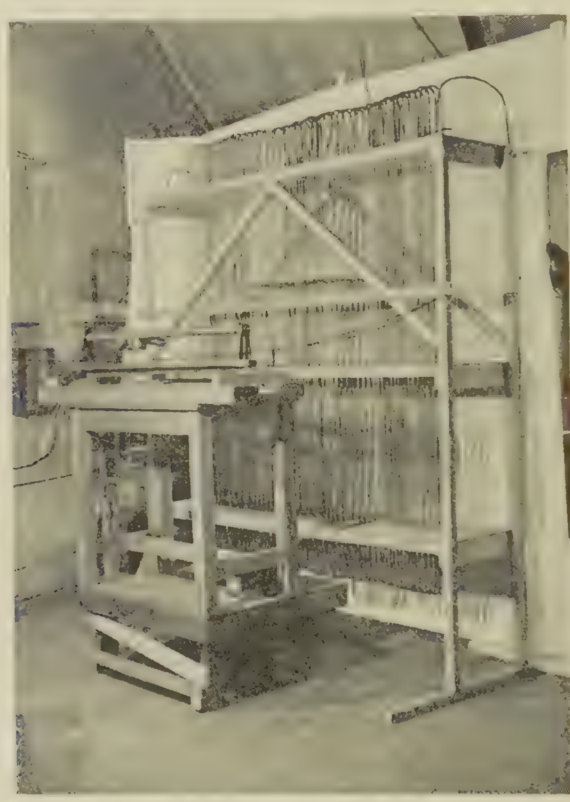


Figure 3.

effective plane of the coil. The adjustment of these mirrors was made by electromagnetic test in a uniform field produced by Helmholtz coils. A horizontal cathetometer CC was mounted with its axis accurately parallel to the central line of the rod system. By reflection of an illuminated scale fixed to the telescope, the search coil could be adjusted with its effective plane either parallel to or normal to the axis of the current system.

A single rod was removed and the electrical connections remade so that the current flowed as before, up in all the rods on one side and down those on the

other side. The search coil was set on the central board in place of the missing rod, carefully levelled, and adjusted so as to be parallel to, or normal to, the horizontal axis of the rod system. A current of some 70 amperes flowing in the rod system was then suddenly reversed. The induced e.m.f. measured the magnetic force normal to the effective plane of the search coil.

In order to be independent of the value of the current in the grid in making comparative measurements at different places, the following method was adopted. The current through the grid was led in series through a pair of Helmholtz coils HH, H′H′, consisting each of a single complete turn, 46 cm. in diameter, of 8 S.W.G. copper, the connection between the two coils being made by wires twisted together non-inductively and so led that they contributed nothing to the field. At the middle of these coils was a circular test coil T, of about 8 cm. diameter, wound with 280 turns of 36 S.W.G. copper wire. This was mounted on a graduated spectrometer table, reading to 1′, so that it could turn about a vertical axis through

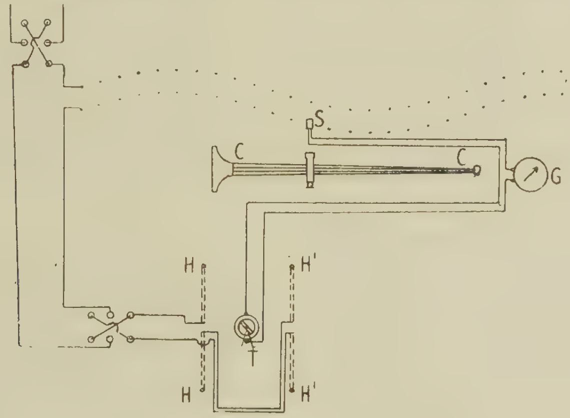


Figure 4.

a diameter, and was connected in series with the search coil and a sensitive galvanometer G. The field at the search coil was measured by setting this test coil at such an angle that, on reversing the current, there was no deflection. The reading was actually made by noting the (very small) galvanometer deflection for a setting near the correct one, and then altering the setting slightly so as to give a galvanometer deflection in the other direction, the setting for zero deflection being obtained by interpolation. The current through the Helmholtz coils was then reversed and another reading taken so as to avoid zero error. The average angle between the effective plane of the test coil and the axis of the Helmholtz coils being θ, $\sin\theta$ measures the magnetic force at the search coil.

For each rod, measurements were made with the search coil in two positions, normal to and parallel to the axis of the grid. In this way the two components of the magnetic force were determined, giving the magnitude, on an arbitrary scale, and direction of the force. Measurements were usually made at the positions of some, or all, of the twenty back rods as well as at the twenty front

rods which go to a wave-length, as a check: the agreement between the values in symmetrically equivalent positions was usually well within 1 per cent.

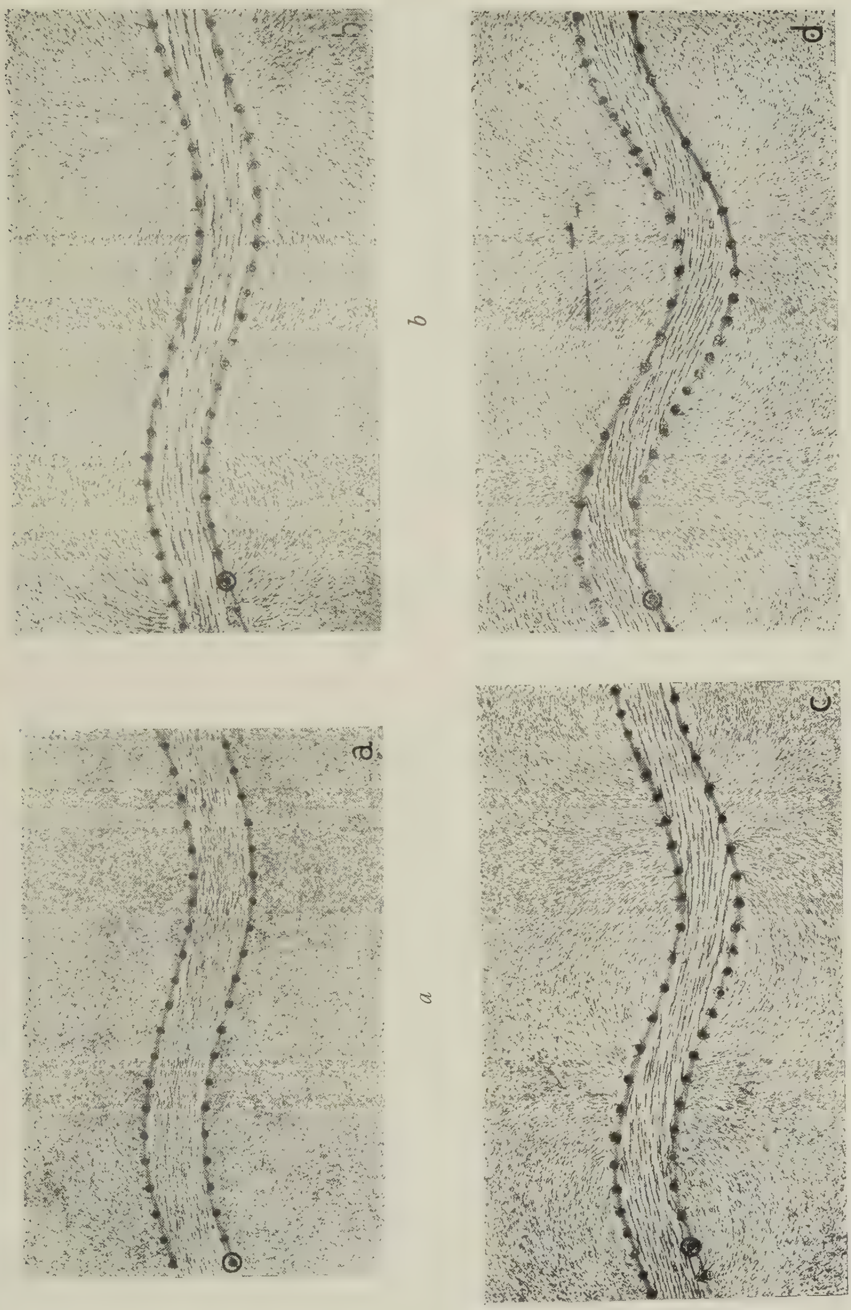
In this way the velocity at any point was found for a given distribution of the rods, and to obtain the next distribution the position of each rod had to be displaced by an amount proportional to the velocity. The scale used was such that the displacement of each rod, in the early stages, was somewhere about 10 per cent of the wave-length. Boards and celluloid were then drilled with holes in the displaced position, and the process repeated with the new set-up. In all eight consecutive dispositions of the current system were obtained.

Since the magnetic force is everywhere in the same direction as the velocity would be for the corresponding distribution of line-vortices, the magnetic lines of force give the lines of flow in the vortex case. These lines of force were obtained with iron filings in the usual way; the filings were then fixed with shellac solution and photographed. In order to obtain them clearly, large currents are necessary; about 250 amperes were used. Since the resistance of the grid is very low, a small voltage suffices. Actually a large Exide accumulator battery, of 100 amp.-h. capacity, was used, which stood up to the demands very well. The lines can be seen in the figures of plates 4 and 5.

The successive positions of the vortex sheet lettered *a* to *g* are shown in plates 4 and 5, in which the vortices are joined up in their order by a continuous line. A striking feature of the progress of the disturbance is the development of the curled back spurs. The one at the crest was expected from the experimental results on the circular jet: it was accompanied, however, by a smaller, but distinct, second spur. This unexpected feature seemed to offer a good opportunity for seeing if there is fair correspondence between the idealized uniform jet of a perfect fluid, with neglect of inertia terms, and the experimental jet. It was therefore decided to investigate the instability of the plane jet experimentally, as will shortly be described.

It will be seen that the lines of flow, as indicated by the magnetic lines of force, do not indicate quite the same boundaries of the jet as does the line joining the vortices. The lines of flow, for instance, tend to show a large detached vortex in place of the main spur, a vortex into which the spur would clearly develop in a slightly later stage. This discrepancy is, no doubt, due to the wide spacing of some of the elementary vortices at the later stages of the development. It is not practical, however, to take a much greater linear density of vortices, and Rosenhead in his mathematical investigation took a lesser density. The difference between lines of flow and vortex sheet is not serious: both indicate the development of a double spur as a stage in the breaking up of the plane jet. The line joining the vortices is probably the best indication of the position taken up by a continuous vortex sheet, and so of the boundary of a real jet.

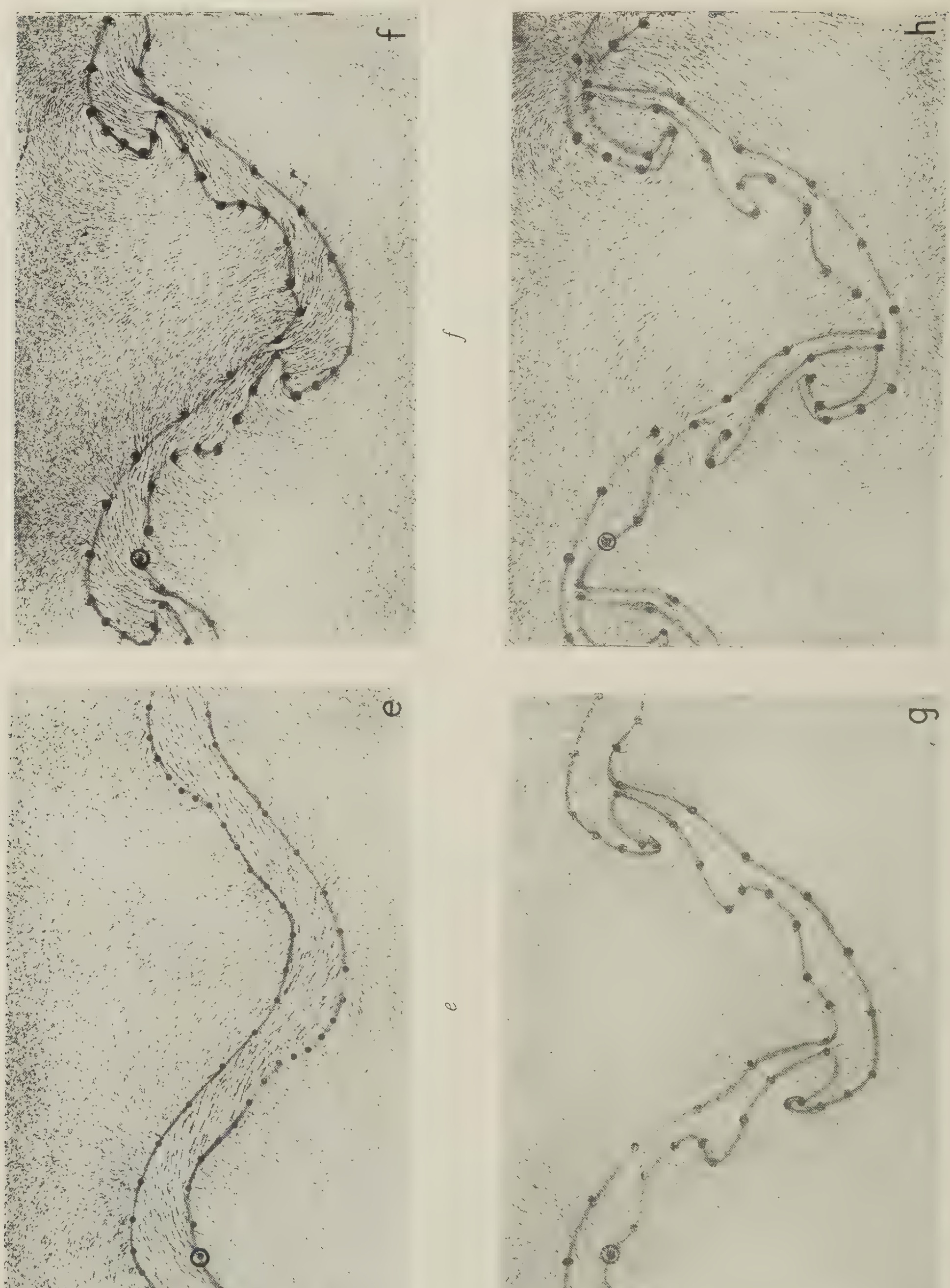
It is suggested that this method of using the hydro-magnetic analogy may be useful as a general method for investigating two-dimensional distributions of vorticity.



a *b* *c* *d*

Plate 4. Development of instability of a plane jet, worked out by the hydro-magnetic analogy.

To face page 344



f *e* *h* *g*

Plate 5. Development of instability of a plane jet, worked out by the hydro-magnetic analogy (*continued*).

§8. THE EXPERIMENTAL INSTABILITY OF THE PLANE JET

To obtain a jet which should approximate to the two-dimensional case represented by the hydro-magnetic model, an orifice $2{\cdot}03 \times 0{\cdot}03$ cm., with a tapering approach, was used, as in my investigation of the velocity distribution in a flat jet. To secure freedom from fluctuations (see Andrade, 1939) the jet issued vertically downwards into clear water in a glass tank, the reservoir supplying the jet being filled with water rendered cloudy by a few drops of milk, a fluid which photographs very well.

This jet was not sensitive to sound. The whole range of frequencies from about 10 to 4000 $\sim$, in intensities much greater than those required to activate the round jet, was applied without result. To produce the disturbance which it was desired to study it was found necessary to cause the nozzle to execute transverse vibrations by mechanical means. In order to obtain a wave-length bearing somewhat the same ratio to the width of jet as that used in the hydro-magnetic analogy, the vibration of the jet must be slow, which has a further advantage already indicated. It was produced by a pendulum, with a heavy adjustable bob, rocking on knife edges fixed to a hollow brass horizontal beam. This beam was attached rigidly to the nozzle, its length being normal to the length of the orifice, and was carried by glass rollers resting on top of the tank. The plane of vibration of the pendulum was parallel to the beam, and the oscillations of the jet were thus transverse to the orifice. The jet was photographed by an electric spark between magnesium spheres, with a lens of 8 in. focal length, f 2·9 aperture.

The amplitude of disturbance could be varied either by maintaining a constant velocity of flow and changing the amplitude of the pendulum, or by keeping the amplitude of the pendulum constant and changing the velocity of flow, i.e. the vorticity at the nozzle. A series of pictures was obtained by each method. In plate 6, *a* shows the undisturbed jet, with an outflow of 1000 cm^3/h., while the other photographs show the disturbance produced with a fixed amplitude of the disturbing pendulum, of period 2·06 sec., and varying velocities of flow. The flow is best expressed in terms of $K = M/\rho$ (where M is the momentum issuing from the orifice per second per unit length and ρ the density of the liquid) and the dimensionless quantity $R = K/v\nu$, where v is the velocity and ν the kinematic viscosity. The following are the values:

Plate 6

Photograph	*a*	*b*	*c*	*d*	*e*	*f*
K	0·637	0·570	0·761	0·804	0·897	0·936
R	13·8	13·0	15·1	15·5	16·4	16·8

Plate 7 shows the effect of varying the amplitude of the pendulum, the period again being 2·06 sec. and the values for the outflow $K = 0{\cdot}700$, $R = 14{\cdot}5$. The values of the amplitude for photographs *a* to *e* were respectively 6·2, 7, 8, 9·3 and 12 cm., which furnish an arbitrary scale.

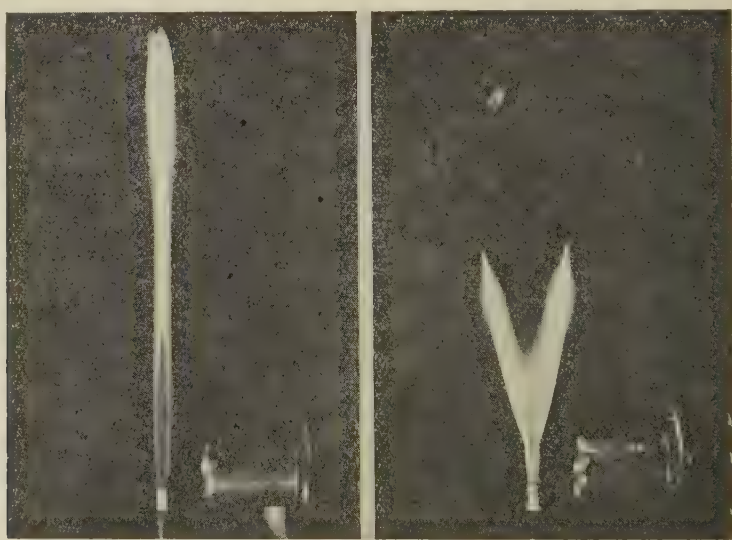
It will be seen that as the disturbance develops, a very sharp thin spur forms

at the crests of the sine wave, closely corresponding to the fine spur shown in the theoretical figures *g* and *h* of plate 5. Not only this, but a secondary spur, developed nearer the nozzle than the main spur, is clearly visible in all the photographs, particularly clearly in some. This corresponds well to the secondary spur shown in *g* and *h*, plate 5. In the early stages of development of the spur a general correspondence can also be traced between the experimental jet and the theoretical jet.

It is considered that the experiments suffice to show that in the case of the two-dimensional jet, the action of the parallel vortex sheets on one another, which can be predicted from the fundamental conception of the behaviour of vortices in perfect fluids, is quite sufficient to explain the manner in which the growing instability of the actual jet develops from a sine wave. From this the conclusion is drawn that the instability of the circular jet can be similarly explained.

§ 9. THE SOURCE OF SENSITIVENESS

We have now dealt with the way in which the disturbance develops, but we still have to discuss the question as to why it develops so easily—why is the sensitive flame sensitive? All experiment points to the orifice as being the source of the



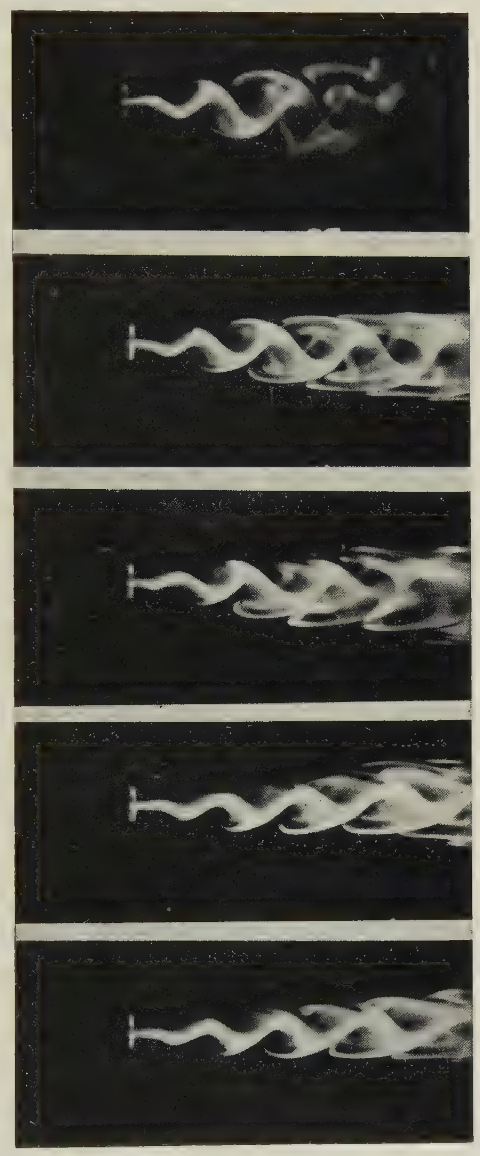
(*a*) (*b*)

Figure 5.

sensitiveness to sound, but experiment further shows that the flat jet is not nearly so sensitive to sound disturbance as the round jet, while the single vortex sheet is not sensitive at all. The growth of a disturbance is, then, clearly due to the influence of one vortex sheet on the opposite parallel, or nearly parallel, sheet. The hydro-magnetic experiments have shown that there is a good first-approximation correspondence between the ideal plane jet and the actual case.

a *b* *c* *d* *e* *f*

Plate 6. *a*, undisturbed jet ; *b* to *f*, disturbed jet, various velocities of outflow.



a *b* *c* *d* *e*

Plate 7. *a* to *e*, disturbed jet, various amplitudes of disturbance.
Experimental instability of a plane jet.

To face page 346

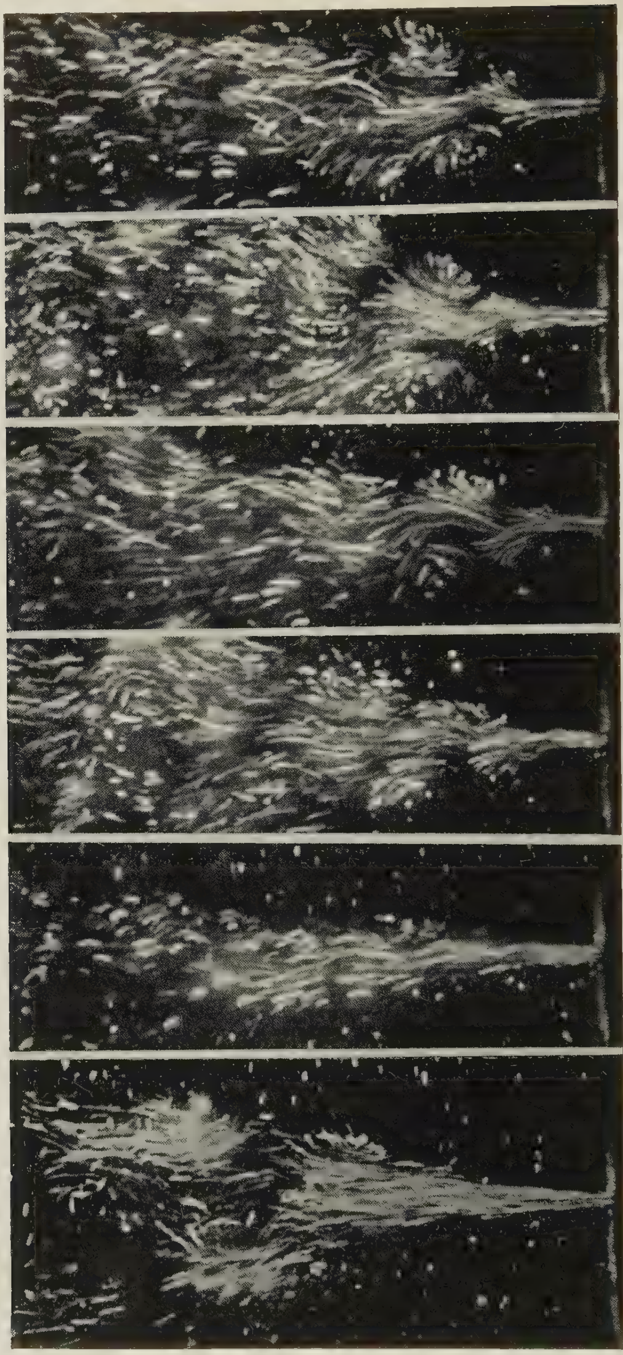


Plate 8. Instability of a water-into-water jet: *a*, due to high velocity; *b* to *f*, due to sound disturbance.

There seems little doubt that the growth of the disturbance in the case of the circular jet is to be attributed to the action of one side of the displaced conical vortex sheet on the other side, the active parts being those in the neighbourhood of a plane through the axis of the jet and the displacement vector. This view accords with the well-known changes of sensitivity for different directions of the sound when the nozzle is slightly asymmetrical, for it would lead us to expect increased sensitivity when the jet is so disposed that the operative parts of the sheet are near together. I have found with an orifice of roughly elliptical shape (axes 2·3 and 1·5 mm.), that the jet is very much more sensitive when the direction of the sound lies along the minor axis than when it lies along the major axis. Photographs (*a*) and (*b*) of figure 5 show the flame from the orifice exposed to the same sound, the direction of propagation being along the major axis in the first case, and along the minor axis in the second case.* Our general conclusion is, then, that there is no need to invoke anything beyond the classical behaviour of vortices and vortex sheets to explain the general method of break-up of the sensitive flame.

On the view put forward here, the sensitiveness is due to the transverse motion of the heavy vorticity at the orifice with reference to the surrounding fluid. With an approximately plane wave of sound, travelling in a direction normal to the axis of the jet, the orifice is the only place where a part of the jet is displaced relative to any other, but by using a small beam of sound, from a small diaphragm with a restricted orifice, as shown in figure 5, and holding it close to the jet, so that the sound travels either normally, or at 45° to the axis, a sinusoidal disturbance may be applied to any part of the jet. Experiments on these lines show that the flame is still sensitive over a range of some centimetres from the origin, but that the sensitiveness increases very markedly as the orifice is approached. The same thing has been shown with the water jet.

To explain the behaviour of the circular jet, and to contrast it with that of the plane jet, it seems advisable to calculate how the vorticity dies off as we recede from the orifice. For the circular non-turbulent jet from a point orifice (Schlichting, 1933) we have

$$u=\frac{3}{8\pi}\frac{K}{\nu x}\frac{1}{(1+\xi^2/4)^2},\quad v=\frac{1}{4}\sqrt{\left(\frac{3K}{\pi}\right)}\cdot\frac{1}{x}\cdot\frac{\xi-\xi^3/4}{(1+\xi^2/4)^2},\qquad \ldots\ldots(1)$$

where
$$\xi=\frac{1}{4}\sqrt{\left(\frac{3K}{\pi}\right)}\cdot\frac{1}{\nu}\cdot\frac{y}{x}\quad\text{and}\quad K=J/\rho,$$

J being the momentum crossing a plane normal to the axis of the jet per second, ν the kinematic viscosity and ρ the density of the fluid. The distance x is measured parallel to the axis, from the point origin of the jet, y is the distance measured normal to this axis, and u and v are the x and y components of velocity.

* See also G. B. Brown, 1932, p. 169.

Mr. Tsien and I have shown that for a jet from a finite orifice of radius a, with slowly converging approach, the distribution is represented by this formula if x is measured from an effective origin represented by

$$x_0 = -0{\cdot}160Ra, \quad \text{where } R = u_0 a/\nu, \qquad \ldots\ldots(2)$$

u_0 being the velocity at the orifice. The agreement with experiment is very good at distances of some 8 diameters of the jet and greater: unfortunately measurements could not be made closer to the jet.

If $P = \dfrac{1}{4\nu}\sqrt{\dfrac{3K}{\pi}}$,

$$\zeta = \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} = -\frac{P^2\nu}{(1+P^2y^2/4x^2)^3}\frac{y}{x^3}\left\{\frac{3P^2y^2}{2x^2} - 2(1-P^2)\right\},$$

$$\frac{\partial\zeta}{\partial y} = \frac{\nu P^2}{x^3(1+P^2y^2/4x^2)^4}\left\{\frac{3}{2}\frac{P^2y^2}{x^2}\left(\frac{3P^2y^2}{2x^2} - 2(1-P^2)\right)\right.$$
$$\left. - \left(1+\frac{P^2y^2}{4x^2}\right)\left(\frac{9P^2y^2}{2x^2} - 2(1-P^2)\right)\right\}.$$

The condition for ζ maximum at x constant reduces to

$$9P^4y^4/x^4 - (P^2y^2/x^2)(56-20P^2) + 16(1-P^2) = 0$$

or, if

$$P^2y^2/x^2 = \alpha,$$

$$9\alpha^2 - (56-20P^2)\alpha + 16(1-P^2) = 0 \qquad \ldots\ldots(3)$$

is the locus of maximum vorticity. That is,

$$y/x = \phi(P),$$

or the locus of maximum vorticity is a cone through the effective origin whose angle is a function of a dimensionless quantity P, given by the momentum, density and viscosity of the jet.

The vorticity ζ may be written

$$\zeta = \frac{1}{x^2} f(P, y/x),$$

or, since y/x is a function of P only along the locus of maximum vorticity, the maximum vorticity is given by

$$\zeta_{\max} = \frac{1}{x^2}\psi(P),$$

where ψ is a complicated function of P whose exact value can be written down if desired. The vorticity, then, dies off inversely as the square of the distance from the effective origin of the jet.

If P is large, it will be found that (3) gives

$$\alpha = 0{\cdot}4.$$

Even if P is 10, this is substantially correct ($\alpha = 0{\cdot}4013$), and for larger values of P the approximation is, of course, closer. In the experiments of Andrade and

Tsien the lowest value of P was about 24, so that this approximation is justifiable for all practical cases.

Hence for all ordinary values of P the locus of maximum vorticity is given by

$$y/x = \sqrt{0{\cdot}4}/P = 0{\cdot}632\ldots/P$$

and
$$\zeta_{\max} = -\frac{\nu}{x^2}\frac{P\alpha^{1/2}}{1{\cdot}1^3}\{\tfrac{3}{2}\alpha - 2 + 2P^2\}$$

$$= -0{\cdot}950\ldots\frac{\nu P^3}{x^2},\ \text{since } P^2 \text{ is very large compared to } 0{\cdot}7,$$

$$= -0{\cdot}0771\ldots\frac{\nu}{x^2}R^3,$$

since $K = \mathcal{J}/\rho = \pi^2 a^2 u_0{}^2 = \pi R^2 \nu^2$, so that

$$P = \frac{\sqrt{3}}{4}R$$

and the maximum vorticity at the orifice is, from (2),

$${}_0\zeta_{\max} = -3{\cdot}01\frac{\nu R}{a^2}.$$

In the case of the plane jet (Bickley, 1937),

$$u = c_1(K^2/\nu x)^{1/3}\,\text{sech}^2\,\xi,$$
$$v = c_2(K\nu/x^2)^{1/3}[2\xi\,\text{sech}^2\,\xi - \tanh\xi],$$

where
$$\xi = c_3(K/\nu^2)^{1/3}yx^{-2/3},\quad K = M/\rho.$$

M = momentum issuing from the orifice per second per unit length of jet normal to the xy plane, and c_1, c_2, c_3 are numerical constants with the respective values $0{\cdot}4543\ldots$, $0{\cdot}5503\ldots$, $0{\cdot}2751\ldots$ This gives

$$\begin{aligned}\zeta = \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} &= \tfrac{2}{3}c_2(K\nu)^{1/3}x^{-5/3}(2\xi\,\text{sech}^2\,\xi - \tanh\xi)\\ &\quad + \tfrac{2}{3}c_2c_3(K^2/\nu)^{1/3}yx^{-7/3}\,\text{sech}^2\,\xi\\ &\quad - \text{sech}^2\,\xi\tanh\xi\{\tfrac{8}{3}c_2c_3(K^2/\nu)^{1/3}yx^{-7/3}\xi + 2c_1c_3(K/\nu)x^{-1}\}\\ &= 2c_2c_3(K^2/\nu)^{1/3}yx^{-7/3}\,\text{sech}^2\,\xi - \tfrac{2}{3}c_2(K\nu)^{1/3}x^{-5/3}\tanh\xi\\ &\quad - c_3(K/\nu)x^{-1}\,\text{sech}^2\,\xi\tanh\xi(\tfrac{8}{3}c_2c_3y^2x^{-2} + 2c_1).\end{aligned}$$

The value of $R = K/\nu\nu$ which corresponds to instability is about 30, and for the slowest jets used in my investigation of the distribution of velocity in the plane jet (Andrade, 1939) R was about 10. The corresponding values of K, with the orifice employed, were 3 and 0·33. In the investigation to which reference has just been made I find, for the virtual origin of the jet, $-x_0 = 0{\cdot}65\,Rw$, where w is the width of the orifice. The smallest value of x that need be considered is x_0. With the values here contemplated it can be shown that, for reasonable values of y (that is, omitting values of y/x which make ζ so small that they do not interest us), the first three terms are small compared to the fourth, or

$$\zeta = -2c_1c_3K/\nu\,x^{-1}\,\text{sech}^2\,\xi\tanh\xi$$
$$= -0{\cdot}250\ldots K/\nu\,x^{-1}\,\text{sech}^2\,\xi\tanh\xi$$

on inserting the values of c_1 and c_2.

Since in the hydro-magnetic investigation of the growth of the instability of the plane jet the vorticity was taken as two sharp sheets, whereas actually there is a continuous distribution of vorticity across the jet, it is of interest to enquire how sharp are the maxima of vorticity. Figure 6 shows vorticity ζ plotted against y for various values of x, K being taken as 1 and ν having the value for water, viz. 0·01. This means $R = 17{\cdot}3$, which is a representative value. The distribution of vorticity is shown for the part of the jet to the left of the axial plane only:

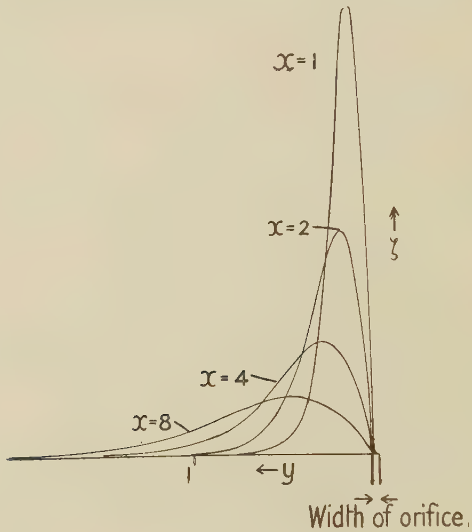


Figure 6.

for the other side of the jet it is the same, with the sign reversed, as indicated in figure 2 (c), which is a sketch of a typical distribution. Figure 2 (d) shows the distribution in the case represented by the hydro-magnetic analogy.

It will be seen that for values of x up to 1, i. e. up to $33w$, the maximum is very sharp, and for higher values of x fairly sharp, until we get to $x = 4$ and 8. The simple vortex sheets are therefore a fair first approximation.

The condition for ζ to be a maximum is

$$\frac{\partial \zeta}{\partial y} = C \operatorname{sech}^2 \xi \{\operatorname{sech}^2 \xi - 2 \tanh^2 \xi\} \frac{\partial \xi}{\partial y} = 0$$

or

$$\tanh \xi = 1/\sqrt{3} = 0{\cdot}577\ldots, \quad \xi = 0{\cdot}658\ldots$$

The locus of maximum vorticity is therefore

$$y = 2{\cdot}392\ldots (K/\nu^2)^{-1/3} x^{2/3}$$

and the maximum vorticity is given by

$$\zeta_{\max} = -0{\cdot}250\ldots \frac{1}{\sqrt{3}} \frac{2}{3} (K/\nu)\, x^{-1}$$

$$= -0{\cdot}0962\ldots (K/\nu)\, x^{-1}.$$

Since the virtual origin of the flat jet is given by

$$-x_0 = 0{\cdot}65Rw,$$

the maximum vorticity at the orifice is

$$_0\zeta_{\max} = -0{\cdot}0962\ldots(K/\nu)(0{\cdot}65Rw)^{-1}$$
$$= -0{\cdot}148\ldots\frac{\nu R}{w^2}.$$

It is realized that these calculations represent only a rough approximation in the immediate neighbourhood of the orifice, since at the orifice itself the actual distribution of velocity must differ somewhat from that assumed by Andrade and Tsien and Andrade. At quite a short distance from the orifice, however, our assumptions are closely obeyed, and even at the orifice itself the calculations give a fair guide.

In the light of these results the hydro-magnetic analogy could be carried out to a better degree of approximation by arranging the current-carrying wires along the locus found for the maximum vorticity and spacing the wires so that the line density corresponded to the law found for the falling off of the maximum vorticity. If someone in America, say, should judge it worth while to do this I should see the results with gratification.

If we now compare the round jet with the flat jet we see, firstly, that for the former the maximum vorticity diminishes inversely as the square of the distance from the effective origin, while for the latter the maximum vorticity diminishes inversely as the first power of the distance from the virtual orifice. Further, if we take the critical value of R for instability as 370 for the former and 30 for the latter (Andrade and Tsien, 1937 ; Andrade, 1939), we have for the maximum vorticity at the origin, when the jet is on the point of instability, $-1114\nu/a^2$ for the round jet and $-4{\cdot}44\nu/w^2$ for the latter. The vorticity at the origin is, then, some 1000 times as great for the round jet as for the flat jet when both jets are on the point of instability, if the diameter of the round jet is equal to the width of the flat jet ($w = 2a$). With the values used in these experiments, $a = 0{\cdot}045$ cm., $w = 0{\cdot}03$ cm., the vorticity ratio is about 100.

The sensitiveness of the round jet as compared with the relative insensitiveness of the flat jet is to be sought in these results. Sound causes the break-up of the jet by the relative transverse movement of the orifice and the surrounding fluid, and the seat of the sensitiveness is in the neighbourhood of the orifice. We may therefore consider the effect of the sound to be a transverse vibration of the vorticity at the orifice, which acts on the vortex sheet in its immediate neighbourhood, the instability developing as the jet travels out in the way traced in detail for the double plane vortex-sheet. Heavy vorticity at the orifice will therefore give sensitiveness.

This, however, is not the whole story. We must enquire how it is that, with the round jet, so heavy a vorticity at the orifice is compatible with stability for the undisturbed jet. The solution lies in the fact that when the velocity of the jet

is increased until turbulence ensues, the break-down is not initiated at the orifice, but at the part of the jet remote from the orifice. An indication of this is that the first signs of instability are a flickering of the remote end of the jet, particularly noticeable in the sensitive flames; but it might be argued that this is because the disturbance, initiated at the orifice, has not grown into a visible movement until it has travelled the whole length of the jet. Stronger evidence is furnished by the photographs of plate 8, taken with a flat water-into-water jet containing particles of a light aluminium-magnesium alloy. *a* represents the jet in an unstable condition owing to high velocity of efflux (800 $cm^3/h.$). *b* to *f* represent the jet at a much lower velocity of efflux (550 $cm^3/h.$), with the orifice constrained to vibrate at 4 cycles/sec., the exposure being 1/16 second. It will be seen that in the case of the instability due to high velocity the jet travels in an unbroken manner for some distance, the large disturbances taking place remote from the orifice. When the instability is produced by vibration of the orifice, the disturbance begins close to the orifice, with disengagement of vortices alternately to either side, the contrast being particularly marked if we compare *a* with, say, *d*, the remote disturbance in the two cases being about the same. Other evidence in this direction is furnished by an experiment described later.

If, then, the sensitiveness to sound has its seat at the orifice, while the instability of the unserenaded flame is initiated by chance disturbances of the far end of the flame, the condition for sensitiveness should be a high vorticity at the orifice, dying off rapidly. This is just what we have in the circular jet, where the vorticity is inversely as x^2. With the flat jet, on the other hand, the vorticity is inversely as x, and the decay of vorticity is far less marked.

These considerations explain an interesting experiment of Zickendraht (1932). He found that the sensitiveness of the flame could be raised either by putting a fine wire just across one side of the circular orifice, or by bevelling the bore of the tube back at the orifice, so as to form a sharp edge past which the issuing gas flowed. Either of these devices will increase the vorticity at the orifice without modifying it at the remoter parts of the jet. The theory here put forward had, in fact, led me to suppose that the sensitiveness could be increased by a sharp edge of the kind just described, and I was about to carry out the experiment when I found that it had already been done.

Incidentally, the work here described shows that Zickendraht's theory of the sensitive flame can scarcely be valid. He supposes that somehow the sound causes volume pulsations inside the flame on both sides, and puts down the forking of the flame to the repulsions between spheres pulsating out of phase, which have been worked out by Bjerknes. Exactly the same forking, however, takes place with water jets, and volume pulsations can hardly be involved here.

With our experimental jets a was 0·045 cm. for the circular jet, while w was 0·030 cm. for the flat jet. The length of the flat jet of water in water was about 20 cm., while the circular jet of water in water could easily be obtained in a stable state with a length of 50 cm.—the length of the trough—and could clearly have

been longer if the length of the trough had permitted. Taking these lengths, we have the following:—

	Vorticity at orifice	Vorticity at far end
Round jet	5500	13·6
Flat jet	50	1·5

It will be seen that whereas, for the round as against the flat jet, the vorticity is 100 times as great at the orifice, it is only about 9 times as great at the far end of the jet, and, as stated, it was clear that a longer round jet could have been obtained. The figure for the far end is extremely rough for the flat jet, as experimentally the flat jet does not remain flat, but tends towards the circular form at great distances.

Returning to the circular jet, we may enquire how the sensitiveness and the range of frequency for which the jet is sensitive depend upon the diameter of the orifice. The vorticity at the orifice, given by ${}_0\zeta_{\max} = -3{\cdot}01\nu R/a^2$, involves the kinematic viscosity of the fluid, the diameter of the orifice and the velocity at the orifice, which are the variables which may be taken as likely to affect the range of frequency favourable for response. R is, of course, dimensionless and the dimensions of ν/a^2 are those of a frequency. Hence it is permissible to take $\nu R/a^2$ as the expression determining the frequency range of response. We can take as an arbitrary measure of this range the lowest frequency at which a disturbance of standard size takes place, the amplitude of the disturbing sound being either constant or a fixed function of the frequency. No quantitative measurements of the disturbance were made, but it is easy to judge by eye a rough standard. Our oscillator, with resistances once set, gave constant intensity at a particular frequency.

With such rough estimates of standard intensity of sound and standard disturbance, it was found that the smaller the orifice of the circular water jet the higher the range of frequency, and further, that the very small orifices were extremely sensitive. This is in qualitative agreement with the formula, since with each orifice the velocity was taken very near that of instability, which means the same value of R in each case.

If a is fixed, and R is fixed by taking a velocity close to that of instability, then the lower limit of the frequency range should be proportional to ν. This was also qualitatively confirmed with different fluids from the same orifice, the limiting frequency increasing in the order water, air, flame-gases, which is the order of increasing kinematic viscosity.

§ 10. SEPARATION OF THE SENSITIVENESS TO SOUND AND THE SENSITIVENESS TO RANDOM DISTURBANCE

If, while the orifice is the source of sensitiveness to sound, the remote part of the flame is the origin of turbulence due to velocity only, then it should be possible to increase the sensitiveness by shielding the upper part of the flame. The best way to do this appeared to be to surround the upper part of the flame by steady hot moving gases, so as to reduce the gradient of velocity and temperature.

Accordingly the upper part of an ordinary sensitive flame, from a circular steatite nozzle, was surrounded by a tubular furnace, which, in the only experiments carried out before the war came, was maintained at about 400° c. It was found that, as anticipated, this rendered the flame insensitive to the ordinary chance disturbances, such as slight draughts. The addition of the furnace made it possible to increase the gas pressure from 26 to 29 cm. of water without flaring, the height of the flame being increased from 61 to 68 cm. The sensitiveness to sound was much increased. As a rough test a Galton whistle was sounded, under standard conditions, in a side room: this produced no effect on the bare flame adjusted to be on the point of flaring, but produced very marked disturbance, of the usual type, in the flame protected by the furnace and adjusted to the higher stable velocity of efflux thus made possible.

The photographs of plate 8 suggest that it may not be necessary to restrict the use of the sensitive jet to the non-turbulent condition, as has always hitherto been done. *a*, plate 8, shows that, even when there is marked turbulence in the upper part of the jet, due to high velocity of efflux, the lower part of the jet shows nothing of the type of disturbance produced by sound. With stroboscopic observation it may easily be possible to trace the effect of sound on a flaring jet, which, on account of the high orifice velocity, should be very sensitive.

Many odd observations on the sensitive flame can probably be traced to the relative behaviour of the jet close to and remote from the orifice. Such questions as whether it is possible to increase the sensitiveness by altering the orifice from a circular to a slightly elliptical shape, on which there are conflicting opinions, can probably be resolved by such considerations. Does slightly stretching the dimensions of a circular orifice parallel to one diameter decrease the velocity of the remote part of the flame if the average velocity of efflux is maintained? This is clearly a complicated question, which may be governed by various minor considerations.

It does appear, however, that an explanation is now available of the fundamental phenomena, and that a method of approach to the various minor attending problems is to hand.

Such are the trifling questions with which I have endeavoured to beguile your time, such the solutions which I have to offer. Weighty problems of cosmology or of the ultimate nature of the forces which rule the physical world have been dealt with by past lecturers and will, no doubt, be handled by those who will follow me. I have already apologized for setting before you fare lighter than that to which you are accustomed, but I will conclude by using in my excuse the words which John Fletcher set down in his dedication of the *Faithful Shepherdess* to Sir Robert Townshend:

"Yet I oft have seen great feasters
Only for the please the pallet
Leave great meat and choose a sallet."

ACKNOWLEDGEMENT

It is a pleasure to acknowledge the invaluable help with the experiments given by my assistant, Mr. L. Walden, G.M., especially in the matter of the photography.

REFERENCES

Andrade, E. N. da C., 1933. *Proc. Roy. Inst.* **27**, 882 ; 1936. *Ibid.* **29**, 320.
Andrade, E. N. da C., 1939. *Proc. Phys. Soc.* **51**, 784.
Andrade, E. N. da C. and Tsien, L. C., 1937. *Proc. Phys. Soc.* **49**, 381.
Bjerknes, V., 1900–2. *Vorlesungen über hydrodynamische Fernkräfte, nach C. A. Bjerknes Theorie.*
Brown, G. B., 1932. *Phil. Mag.* **13**, 161 ; 1935. *Proc. Phys. Soc.* **47**, 703.
Constable, J. E. R., 1936. *Proc. Phys. Soc.* **48**, 919.
Humby, S. R., 1927. *Proc. Phys. Soc.* **39**, 435.
Rayleigh, Lord, 1896. *Sound*, vol. II, chapter XXI.
Rosenhead, L., 1931. *Proc. Roy. Soc.* A, **134**, 170.
Schlichting, L., 1923. *Z. angew. Math. Mech.* **13**, 260.
Tyndall, J., 1875. *Sound*, 3rd Edition, p. 227.
Zickendraht, H., 1932. *Helv. phys. Acta*, **5**, 317 ; 1934. *Ibid.* **7**, 773.

THE λ3105 BAND OF OD

By M. ISHAQ, M.A., Ph.D.,
Muslim University, Aligarh

Communicated by Dr. R. W. B. Pearse, 20 January 1941

ABSTRACT. The λ3105 band of OD has been photographed in the 4th order of a 10-foot concave grating. The band, like that of OH at λ3122, is the (1, 1) band of a $^2\Sigma^+ \rightarrow {^2\Pi_{\mathrm{inv.}}}$ transition. An analysis of the rotational structure has been made, and the values of the rotational constants have been determined and compared with those for the corresponding OH band.

§1. INTRODUCTORY

In 1934, R. W. Shaw and R. C. Gibbs reported that they had photographed the (0, 0) and (1, 1) bands of OD, but their analysis has not, so far, appeared. Later, the author published the rotational analyses of the (2, 0) and (3, 1) bands (Ishaq, 1937 a) and the (0, 0) band of OD (Ishaq, 1937 b), and also attempted the analysis of the (1, 1) band. This band, being mixed up with the tail end of the (0, 0) band, proved rather difficult, but the analysis has now been successfully carried out, its correctness being shown by the combination differences.

§2. EXPERIMENTAL

The type of discharge tube and the method of excitation were similar to those previously described (Ishaq, 1937). The heavy water was obtained from Imperial Chemical Industries, Ltd., and was 99·2 per cent pure.

The band was photographed in the 4th order of a 10-foot concave grating at the Imperial College, London. The dispersion was about 1·3 A./mm., and

the time of exposure was about four hours. Wave-length measurements were made by comparison with the iron arc, the iron standards being taken from the *Transactions of the International Astronomical Union*, **4**, 1933, supplemented by Kayser's *Handbuch der Spectroskopie*, vol. VII. 313 rotational structure lines have been measured. In the interests of economy their wave-numbers are not given here, but a catalogue similar to that given by the author in 1937 for the (0, 0) and (1, 1) bands of OD has been placed in the archives of the Physical Society.*

§ 3. ROTATIONAL STRUCTURE AND TERM DIFFERENCES

A photograph of the (0, 0) and (1, 1) bands, with their analysis into branches, is reproduced in the plate, figure 1. The structure is similar to that of OH bands, which are attributed to a $^2\Sigma^+ \rightarrow {}^2\Pi_{\text{inv.}}$ transition. An energy-level diagram showingthe branches identified is given in figure 2.

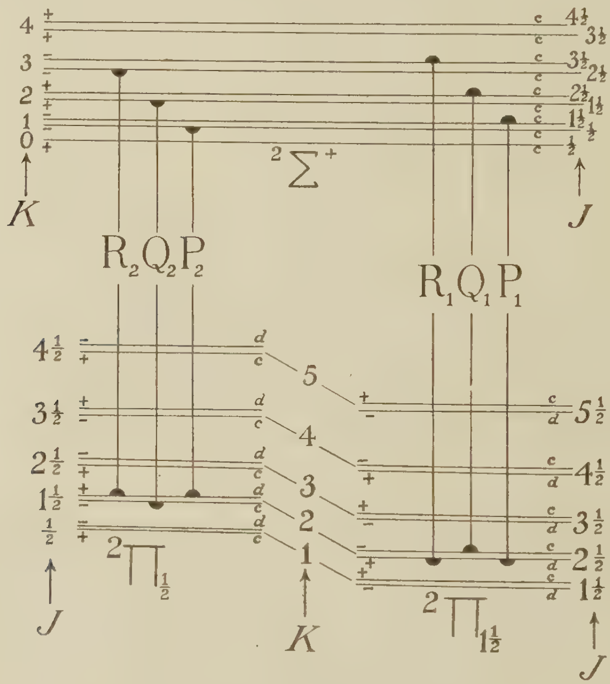


Figure 2. Energy level diagram for a $^2\Sigma^+ \rightarrow {}^2\Pi_{\text{inv.}}$ transition showing the branches identified for the (1, 1) band of OD.

Rotational term differences for the $^2\Sigma$ state

The rotational levels of the $^2\Sigma$ state are treated as singlets, as only the main branches are observed and not the satellites corresponding to the spin-doubling of these levels.

* Three copies are available in the Library of the Physical Society, 1 Lowther Gardens, Exhibition Road, London, S.W. 7, and three have been sent to the Office of the American Institute of Physics, 175 Fifth Avenue, New York. They may be consulted by applying to either of these addresses.—EDITOR.

Figure 1. The 0, 0 and 1, 1 bands of OD. (10-foot concave grating, 4th order.)

To face page 356

The rotational term differences,

$$\Delta_2 F'(K) = F'(K+1) - F'(K-1),$$

are obtained from the combinations

$$\Delta_2 F'(K) = R_1(K+1) - P_1(K+1) = R_2(K+1) - P_2(K+1).$$

The values of these intervals are given in table 1.

Table 1. Rotational term differences for the $^2\Sigma$ state

K	$R_1(K+1)-P_1(K+1)$	$R_2(K+1)-P_2(K+1)$	K	$R_1(K+1)-P_1(K+1)$	$R_2(K+1)-P_2(K+1)$
1	86·95	86·88	11	427·29	427·15
2	121·87	121·11	12	459·96	459·52
3	156·52	156·33	13	491·83	491·69
4	190·98	190·55	14	523·83	523·65
5	225·54	224·73	15	555·16	555·53
6	260·00	259·63	16	586·38	586·32
7	293·66	293·08	17	616·74	616·48
8	327·70	326·55	18	646·92	646·79
9	361·50	360·31	19	676·74	676·24
10	394·52	393·18	20	705·94	

Rotational term differences for the $^2\Pi$ state

In the $^2\Pi$ state each of the sub-states $^2\Pi_{1\frac{1}{2}}$ and $^2\Pi_{\frac{1}{2}}$ has a set of rotational levels which are split up into two close sub-levels on account of Λ-doubling.

The rotational term differences,

$$\Delta_1 F''(K+\tfrac{1}{2}) = F''(K+1) - F''(K),$$

are evaluated from the following combinations:

$$\text{for } {}^2\Pi_{1\frac{1}{2}} \begin{cases} \Delta_1 F_{1cd}''(K+\tfrac{1}{2}) = R_1(K) - Q_1(K+1), \\ \Delta_1 F_{1dc}''(K+\tfrac{1}{2}) = Q_1(K) - P_1(K+1)\ ; \end{cases}$$

$$\text{for } {}^2\Pi_{\frac{1}{2}} \begin{cases} \Delta_1 F_{2cd}''(K+\tfrac{1}{2}) = R_2(K) - Q_2(K+1), \\ \Delta_1 F_{2dc}''(K+\tfrac{1}{2}) = Q_2(K) - P_2(K+1). \end{cases}$$

Their values are given in tables 2 and 3. The values for the (3, 1) band, λ2756, of OD are taken from the analysis by Ishaq (1937 a) except those derived from the following members of the Q_2 branch for which new wave-numbers have been obtained:

	Revised value	Previous value
$Q_2(8)$	36044·1	36046·2
$Q_2(12)$	35923·1	35925·7
$Q_2(13)$	35884·0	35886·2
$Q_2(14)$	35841·2	35844·8

The differences between the values of $\Delta_1 F_{dc}'(K+\frac{1}{2})$ and $\Delta_1 F_{cd}''(K+\frac{1}{2})$ are a measure of the Λ-doubling, for

$$\begin{aligned}\Delta_1 F_{dc}''(K+\tfrac{1}{2}) - \Delta_1 F_{cd}''(K+\tfrac{1}{2}) &= [F_d''(K+1) - F_c''(K+1)] + [F_d''(K) - F_c''(K)] \\ &= \Delta\nu_{dc}(K+1) + \Delta\nu_{dc}(K) \\ &= 2\delta_{dc}(K+\tfrac{1}{2}), \text{ say.}\end{aligned}$$

The quantity $\delta_{dc}(K+\frac{1}{2})$ is thus the mean of the Λ-doubling in the levels $F''(K+1)$ and $F''(K)$. Its values are shown in table 4.

§ 4. CALCULATION OF THE ROTATIONAL CONSTANTS

(1) The term differences for the $^2\Sigma$ state are well represented by the energy function

$$F'(K) = B_v'K(K+1) + D_v'K^2(K+1)^2 + \ldots$$

with values of B_v' and D_v' given in table 5. The differences are given by

$$\Delta_2 F'(K) = (4K+2)[B_v' + 2D_v'(K^2+K+1) + \ldots]$$

In table 5 the values for OH are due to Tanaka and Koana (1933) and those for OD to Ishaq (1937 a and b).

(2) For the calculation of the rotational constants of the $^2\Pi$ state, the closed formula due to Hill and Van Vleck (1928) has been used:

$$F(J) = B\left[(J+\tfrac{1}{2})^2 - \Lambda^2 \pm \tfrac{1}{2}\left\{4(J+\tfrac{1}{2})^2 + \frac{A}{B}\left(\frac{A}{B} - 4\right)\Lambda^2\right\}^{\frac{1}{2}}\right] + \ldots.$$

For a Π state, $\Lambda = 1$.

Table 2. Rotational term differences for $^2\Pi_{1\frac{1}{2}}$ sub-state

K	$\Delta_1 F_{1cd}''(K+\frac{1}{2}) = R_1(K) - Q_1(K+1)$ (1, 1) band, λ3105	(3, 1) band, λ2756	K	$\Delta_1 F_{1dc}''(K+\frac{1}{2}) = Q_1(K) - P_1(K+1)$ (1, 1) band, λ3105	(3, 1) band, λ2756
1	45·27	45·1	1	44·99	45·0
2	63·14	63·2	2	63·22	63·7
3	81·64	81·7	3	81·45	81·4
4	100·14	99·9	4	99·65	99·8
5	118·63	118·5	5	117·91	118·3
6	137·28	137·3	6	136·27	136·2
7	156·03	155·7	7	154·49	154·2
8	174·30	174·1	8	172·87	172·4
9	192·91	193·0	9	191·05	190·6
10	211·66	211·3	10	209·23	208·7
11	229·98	230·2	11	227·17	226·8
12	248·07	248·2	12	244·79	244·2
13	266·32	266·3	13	261·87	261·8
14	284·66	285·1	14	279·39	279·9
15	302·93	302·4	15	296·59	
16	320·49		16	313·64	

Table 3. Rotational term differences for the $^2\Pi_{\frac{1}{2}}$ sub-state

K	$\Delta_1F_{2cd}''(K+\frac{1}{2})=R_2(K)-Q_2(K+1)$ (1, 1) band, λ3105	(3, 1) band, λ2756	K	$\Delta_1F_{2dc}''(K+\frac{1}{2})=Q_2(K)-P_2(K+1)$ (1, 1) band, λ3105	(3, 1) band, λ2756
1	28·92	33·0	1	28·84	29·9
2	50·58	49·8	2	50·46	48·9
3	70·65	69·6	3	71·06	71·2
4	90·33	89·8	4	90·65	91·1
5	110·20	109·7	5	110·55	110·9
6	129·80	129·5	6	130·61	130·4
7	149·19	149·6	7	149·05	150·1
8	168·66	168·6	8	168·13	168·3
9	187·63	187·8	9	187·07	187·9
10	206·69	206·6	10	205·55	206·7
11	225·35	225·2	11	225·04	224·8
12	244·11	245·3	12	242·99	243·0
13	262·56	263·0	13	260·91	260·8
14	280·77		14	278·87	279·1
15	299·25		15	296·97	
16	316·57		16	313·49	
17	334·84		17	331·18	
18	352·08		18	347·82	
19	369·00		19	364·35	
20	385·95				
21	402·53				
22	419·33				
23	436·24				
24	452·14				
25	467·94				

Table 4. Λ-doubling in the $^2\Pi$ state

K	$^2\Pi_{1\frac{1}{2}}$	$^2\Pi_{\frac{1}{2}}$	K	$^2\Pi_{1\frac{1}{2}}$	$^2\Pi_{\frac{1}{2}}$
1	−0·14	−0·04	11	−1·40	−0·16
2	+0·04	−0·06	12	−1·64	−0·56
3	−0·10	+0·20	13	−2·22	−0·82
4	−0·25	+0·16	14	−2·63	−0·95
5	−0·36	+0·18	15	−3·17	−1·14
6	−0·50	+0·40	16	−3·42	−1·53
7	−0·77	−0·07	17		−1·83
8	−0·72	−0·26	18		−2·13
9	−0·93	−0·28	19		−2·33
10	−1·22	−0·57			

Table 5. Rotational constants of OH and OD

v', v''		λ	B_v'	D_v'	B_v''	D_v''
0, 0	OD	3065	9·020	$-0{\cdot}580\times10^{-3}$	9·870	$-0{\cdot}400\times10^{-3}$
	OH	3064	16·955	$-2{\cdot}00\times10^{-3}$	18·516	$-1{\cdot}82\times10^{-3}$
2, 0	OD	2708	8·367	$-0{\cdot}577\times10^{-3}$	9·865	$-0{\cdot}430\times10^{-3}$
	OH	2608	15·283	$-2{\cdot}00\times10^{-3}$	18·524	$-1{\cdot}82\times10^{-3}$
1, 1	OD	3105	8·695	$-0{\cdot}580\times10^{-3}$	9·584	$-0{\cdot}350\times10^{-3}$
	OH	3122	16·117	$-1{\cdot}98\times10^{-3}$	17·809	$-1{\cdot}78\times10^{-3}$
3, 1	OD	2756	8·042	$-0{\cdot}555\times10^{-3}$	9·540	$-0{\cdot}350\times10^{-3}$
	OH	2677	14·458	$-2{\cdot}09\times10^{-3}$	—	—

Splitting the formula into two parts, we get

$$F_1(J)=B\left[(J+\tfrac{1}{2})^2-1+\tfrac{1}{2}\left\{4(J+\tfrac{1}{2})^2+\frac{A}{B}\left(\frac{A}{B}-4\right)\right\}^{\frac{1}{2}}\right]+\ldots.$$

$$F_2(J)=B\left[(J+\tfrac{1}{2})^2-1-\tfrac{1}{2}\left\{4(J+\tfrac{1}{2})^2+\frac{A}{B}\left(\frac{A}{B}-4\right)\right\}^{\frac{1}{2}}\right]+\ldots.$$

On summation and adding a term containing D,

$$F_1(J)+F_2(J)=2B[(J+\tfrac{1}{2})^2-1]+2D[(J+\tfrac{1}{2})^2-1]^2+\ldots,$$

and taking the first differences

$$\Delta_1F_1(J)+\Delta_1F_2(J)=4(J+1)[B+D(2J^2+4J+\tfrac{1}{2})\ldots]$$

The rotational constants B'' and D'' shown in table 5 are obtained by means of this equation.

(3) For the calculation of A, we take the difference of $F_1(J)$ and $F_2(J)$:

$$F_1(J)-F_2(J)=B\left\{4(J+\tfrac{1}{2})^2+\frac{A}{B}\left(\frac{A}{B}-4\right)\right\}^{\frac{1}{2}}.$$

We have therefore

$$\left(\frac{A}{B}-2\right)^2=\{F_1(J)-F_2(J)\}^2/B^2-4\{(J+\tfrac{1}{2})^2-1\}\ldots.$$

If the small effects of Λ-doubling in the $^2\Pi$ state are neglected, the intervals $F_1(J)-F_2(J)$ are given by the combinations

$$F_1(J)-F_2(J)=R_1(J)-Q_2(J)=Q_1(J)-P_2(J).$$

The values of A so derived are given in Table 6.

Table 6. Values of A (cm.$^{-1}$)

J....	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$4\frac{1}{2}$	$5\frac{1}{2}$	$6\frac{1}{2}$	$7\frac{1}{2}$	$8\frac{1}{2}$	$9\frac{1}{2}$	$10\frac{1}{2}$
A....	−139·4	−139·2	−139·8	−139·8	−139·7	−140·2	−139·7	−140·8	−139·2	−138·6

Mean: $A=-139{\cdot}6$

I am glad to take this opportunity of thanking Professor H. Dingle and Dr. R. W. B. Pearse for their continued interest in my work.

REFERENCES

Hill, E. L. and Van Vleck, J. H., 1928. *Phys. Rev.* **32**, 261.
Ishaq, M. 1937 a. *Proc. Roy. Soc.* A, **159**, 110.
Ishaq, M., 1937 b. *Proc. Nat. Inst., India,* **111**, 389.
Shaw, R. W., and Gibbs, R. C., 1934. *Phys. Rev.* **45**, 124.
Tanaka, T. and Koana, Z., 1933. *Proc. Phys.-Math. Soc. Japan,* **15**, 272.

ANEMOMETRY: A CRITICAL AND HISTORICAL SURVEY

By P. A. SHEPPARD,
Imperial College of Science and Technology

Lecture delivered 14 *February* 1941 ; *MS. received* 16 *April* 1941

§1. INTRODUCTION

The measurement of wind had its scientific birth in Italy, in the late 16th century, and in this country sometime prior to 1667, when Robert Hooke described an anemometer to the Royal Society. Since that time the subject has progressively expanded down to the present time. Now, in an age in which aerodynamics and meteorology have assumed much enhanced practical importance, anemometry continues to improve upon its past constructions.

It is not the purpose of the present article to give a detailed historical treatment of the progress in anemometry, but to show the way in which physical principles have been successively applied in the development of apparatus, with notes on the characteristic features of the more important instruments in the several classes of anemometer.

Wind may be natural, the motion of the free atmosphere, or artificially produced as in the flow of air or other gases through pipes, ducts, wind tunnels, etc. Whichever be concerned, the same physical principles are applied in the measurement, though, on account of the generally different features of the motion in the two cases, instruments for measuring the one have developed along rather different lines from those for measuring the other. Some attention will be given in the present article to the measurement of each kind of air motion. But before proceeding to a survey of the instrumental methods employed in wind measurement it is desirable to have in mind the salient features of wind, whether in a laboratory or in the atmosphere. For to the question " What is there to be measured ? " there are several answers, and instrumental needs will vary with the answer given.

§2. CHARACTERISTICS OF AIR MOTION

Wind is a vector quantity and is specified by its speed and direction. In the laboratory, in pipes, tunnels, etc., the general direction of flow is predetermined, but in the atmosphere it is a quantity for measurement.

Air flow may be of two kinds, so-called laminar or streamline flow and turbulent flow, of which the latter is of the greater practical importance both in pipes, etc., and in the atmosphere. The nature of the flow at any instant may be demonstrated by mapping the field with so-called streamlines, at any point of which the tangent gives the direction of flow. In laminar flow, the external conditions remaining constant, the motion is steady, the streamlines remain unchanged, and all fluid elements successively entering the field of flow at a given point traverse the same path, which in such a case is a streamline. For our present purpose the important feature of laminar flow is that, molecular motions excepted, the velocity at any point of the field is not a function of time.

In turbulent flow it is not normally possible to trace out the streamlines, which vary from moment to moment, and fluid elements entering the field successively at a given point do not traverse the same path. The motion is irregular and the fluid velocity at any point varies in speed and direction from one instant to the next. Nevertheless, the motion at any point may be characterized by a mean value, if the period over which the mean is taken is sufficiently long to include a representative set of fluctuations. Alternatively, the mean value may be defined spatially instead of temporally. Every physical instrument does in fact give some kind of mean value over a certain volume determined by its size and over a time depending on its inertia. When the flow is turbulent it is frequently necessary to know other quantities than the mean velocity. Such are:—

(i) Instantaneous departures from the mean value, frequently expressed as turbulent components of velocity u', v', w', along three co-ordinate directions, the mean value of each being zero.

(ii) The arithmetical mean value or root-mean-square value of the turbulent components u', v', w'.

(iii) The frequency distribution of the turbulent components in magnitude and period.

Obviously any one instrument is not likely to be suited to a measurement of mean velocity and also of the above quantities.

In turbulent flow through pipes, wind tunnels, etc., the instantaneous variations from the mean speed are generally not more than a few per cent of the latter, and the frequencies of the fluctuations are mainly high, extending in any case over a considerable range, up to at least some few thousand per second. The turbulence is said to be fine-grained. In the atmosphere, on the other hand, near the ground, the instantaneous variations may be a much greater fraction of the mean speed (see figure 7) and much larger periodicities are involved, of the

order of a second to an hour. To what extent fine-grained turbulence is also present in the atmosphere remains to be investigated.

Finally, the mean velocity at any point in a fluid stream is a function of the distance from the boundary encompassing the flow, so that observations at a single distance from the boundary will be of limited significance if the variation with distance is not also determined. This variation depends on the nature of the flow, * whether laminar or turbulent, but in all practically important cases becomes less as the distance increases.

§ 3. CLASSIFICATION OF ANEMOMETERS

Anemometers may be classified according to the physical principles employed in their design, as follows :—

(1) Time-displacement anemometers, which measure the rate of displacement of some entity of negligible inertia released into the air stream. Such instruments give a direct measure of the velocity " following the motion ".

(2) Velocity meters, in the main rotation anemometers, whose indications are dependent, to a good approximation at least, only on the local velocity of the air stream and not on such properties as air density. In this class are included instruments of the windmill type, rotating cup anemometers and flame air-meters.

(3) Pressure anemometers, which are of two types, depending

(i) on the relation between hydrostatic pressure and velocity at any point in the moving fluid (Bernouilli's equation) ;

(ii) on the pressure exerted by the fluid stream on a suitably mounted plate or other solid body.

In each case the instrumental response depends on the air density as well as on the speed.

Instruments of the first sub-class may be conveniently referred to as manometric anemometers, of which the Pitot-static tube is the most important member, since, suitably designed, it occupies the position of a standard instrument. A modification of the Pitot-static tube for meteorological use is established as a pre-eminent instrument in this and some other countries. Also in the sub-class are instruments whose action depends on a constriction in the flow, e.g. the Venturi tube, plate orifice and shaped nozzle.

The second sub-class comprises instruments of the so-called pressure-plate type.

(4) Anemometers depending on the dissipation of heat or matter from a body placed in the air stream. The most important of the heat-dissipation

* The properties of the flow about a body or through a tube, etc., are determined to an important extent by the Reynolds number vd/ν, where v is the fluid velocity, d a length representative of the body, tube, etc., and ν the kinematic viscosity of the fluid (equal to the coefficient of molecular viscosity divided by the density). There is no fixed value of the Reynolds number at which laminar flow becomes turbulent, but in flow through a pipe the motion will be laminar if vd/ν (d= tube diameter) is less than about 2000.

instruments is the electrical hot-wire anemometer. There is also the Kata-thermometer. Of little more than historical importance is an instrument which depends on the rate of evaporation from a wetted surface as a function of wind speed (Brewster, 1830).

(5) Acoustic anemometers, in which a note is produced whose pitch depends on wind speed. This is a class of no practical importance.

§4. TIME-DISPLACEMENT ANEMOMETERS

Nature presents us with the rudiments of the displacement anemometer in thistledown and suchlike entities which have a very low terminal velocity of fall through the air and which move for all practical purposes with the fluid element into which they are released. An absolute measure of velocity is obtained by timing the motion of such entities.

The first recorded application of the method, some time before 1680, is by Mariotte (1717). Even today the method is not to be despised for the measurement of light winds in the atmosphere. Refinements of the method are seen in the measurement of upper winds in the atmosphere by means of free-lift pilot balloons followed by theodolite, or by balloons carrying wireless transmitting equipment and followed by a suitable ground receiver.

The method has also been used in the laboratory, particularly in investigations of turbulence, by releasing " hot spots " of air into the stream and following their motion by *Schlieren* photography (Townend, 1934), or again by the observation of ultramicroscopic particles suspended in a fluid stream (Fage and Townend, 1932 ; Fage, 1936 a), the particles being observed through a rotating microscope objective.

§5. VELOCITY METERS

The most important instruments of the velocity-meter class are rotation anemometers in which a wind-sensitive rotor turns at a rate determined by the wind-speed. Because of friction at the bearings the rate of rotation does in fact depend very slightly on air density also, but the corrections for its variation are usually negligible. Rotation instruments have a history dating from the mid-eighteenth century, though only during the last century (with the introduction of the rotating cup anemometer and windmill-type airmeter) did they come into wide use.

§6. ROTATING CUP ANEMOMETERS

The rotating cup anemometer arose from experiments by Borda (1763) and Edgeworth (1783) on the resistance of different forms of surface held in a stream of fluid. In particular, the resistance of a cup whose concave side faces the wind is greater than when the convex side faces the wind. Robinson (1846) of Armagh Observatory developed an instrument for meteorological use, employing hemispherical cups, 3 in. diameter, on two crossed horizontal arms, 10·7 in. length

from cup centre to cup centre, mounted on a vertical spindle. Cacciatore (1842) had previously built an instrument for use at Palermo Observatory using four curved sails instead of cups.

The very extensive literature of cup anemometry has been much concerned with the so-called cup factor, i.e. the ratio of wind-speed to cup-speed. Robinson at first erroneously stated this factor to be a constant, independent of cup diameter and arm length, and equal to 3·0. Later work showed that the factor may lie between 2 and 3, depending in an unresolved way on the various structural variables, and that in no case is it a constant, but a function of wind-speed. To overcome the latter feature in some measure it was the standard practice to apply a factor to the records of instruments at meteorological observatories, etc.,

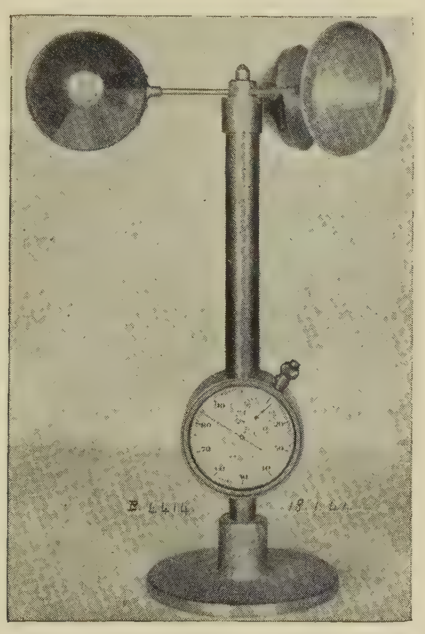


Figure 1. Sensitive portable cup anemometer.

which should be correct at the most probable wind-speed experienced at land stations. This procedure involved errors which, at lower and higher wind-speeds, were frequently considerable.

Investigations in the last two decades have added much to our knowledge of the effect of certain of the variables affecting cup-anemometer characteristics, but a complete mathematical treatment of the motion has not so far proved possible. Of experimental work the most important in the period is due to Patterson (1926) who showed that:

(i) The cup factor was more nearly constant for a 3-cup system than for a 2, 4, 5 or 6-cup system.

(ii) The mean torque on one cup for one revolution was slightly greater for the 3-cup than for the 4-cup system and the torque was less variable with angle of wind incidence in the former case.

(iii) Approach to constancy in the cup factor is best attained by short arms, but this result may not be of general validity.

Patterson's investigations led to the adoption in 1924 of a 3-cup system as standard in the United States and Canada, where the instrument is extensively used.

A further advance in design was made by the discovery by Dryden (1934) that a beaded edge on anemometer cups resulted in the anemometer calibration being less variable with variation in wind-stream turbulence of fine grain than was previously the case.

On the theoretical side Spilhaus (1934) has given a treatment based on the principle of dynamical similitude, in which he has shown the importance, and in some cases the effects, of the following non-dimensional parameters :—

(i) The ratio, d/D, of cup diameter to diameter of cup-centre orbit.

(ii) A Reynolds number $v_0 d/\nu$ (v_0 = value of wind-speed at which cups just cease to rotate, ν = kinematic viscosity of the air).

(iii) The ratio λ/d, where λ is some length characterizing the magnitude of the air-stream turbulence.

(iv) The ratio $\delta/d(D/d-1)$, where δ = diameter of cup arms. (The ratio is representative of the area of the arms exposed to the wind to the area of a cup face).

In particular he has shown that it should be possible to design a cup anemometer with the calibration equation

$$v = v_0 + cu, \qquad \ldots\ldots(1)$$

where v = wind speed,

u = speed of cup centres,

and c = constant,

so that the true wind-speed may be derived from an indicated speed, based on the value of c, by the addition of a simple constant. Such an instrument (figure 1) has been designed by Sheppard (1939). It employs three semi-conical cups, and the anemometer spindle is mounted on miniature ball bearings. The rotation-counting mechanism is a watch movement, the escape lever of which is operated by the rotation of the anemometer spindle. This instrument responds to winds down to about 20 cm./sec. More usually cup anemometers have not been sensitive to winds below about 1 m./sec.

On account of aerodynamic and inertia effects, a cup anemometer over-estimates the mean wind-speed if the wind is gusty. Chree (1895) and Schrenk (1929) have estimated the effect for hemispherical cup anemometers. For those of normal dimensions—9-in. cups on 24-in. arms in a design due to Beckley (1856)

were used for many years at British observatories, whilst semi-portable patterns with 3-in. cups and $7\frac{1}{2}$-in. arms have also been widely used—the over-estimation in a natural wind may amount to as much as 10 or 20 per cent in low winds, decreasing with increase in wind-speed. In Sheppard's instrument the effect has been much reduced.

Various methods, mechanical and electrical, are available for recording wind-speed by means of cup anemometers. In most cases the recorder does in fact give the run of wind from which the mean wind-speed can be deduced over any desired interval of time. The instrument is not normally suited to display the structure of wind, i.e. the succession of gusts and lulls of which it is usually composed, though Maurer (1883) used a governor device to indicate "instantaneous" wind-speed, and in another case (Johnson and Heywood, 1938) the anemometer spindle was linked with a dynamo and recording voltmeter.

The chief merit of the cup anemometer is that its indications are independent of the wind direction in the plane of the cup motion, so that no wind vane is required for meteorological use. Because of its limitations in other respects it has largely fallen out of use as a recording instrument in this country, but continues useful where a portable instrument is required to give mean wind-speed. It has been little used outside the meteorological field, but small sensitive patterns of the Sheppard type may well have uses in wind-tunnel work, etc.

§7. WINDMILL ANEMOMETERS (VANE AIR-METERS)

This type of anemometer (figure 2) consists of a set of light vanes mounted on radial arms attached to a spindle which rotates in jewelled bearings, the spindle being linked with a rotation-counting mechanism. The spindle should lie in the direction of the air stream, so that for meteorological use the instrument needs to be mounted on a wind vane.

Schober (1752) and Woltmann (1790) first developed the vane air-meter on the continent, whilst in this country it is associated with the name of Whewell (1837). Considering the wide use of windmills in earlier times, it is surprising that the instrument was not developed sooner. Various designs of vane air-meter are used, differing in the location of the counter with respect to the vanes, or in the use of curved instead of flat plate vanes; and Ower (1926 a), by reducing the gearing and counting mechanism to the minimum possible, has designed a particularly sensitive pattern suitable for use in light winds down to about 20 cm./sec.

If bearing friction be neglected, the vane speed, u, in a wind of speed v will be such that the wind-speed relative to the vanes is along the latter, i.e., there is no resultant force perpendicular to the vanes. If the vanes at rest are set at an angle θ to the wind direction, then

$$u = v \tan\theta, \qquad \ldots\ldots(2)$$

giving a linear calibration. θ is usually about 45°, so that u and v are approximately equal. This is in contrast to the cup anemometer, in which cup-speed

is much less than wind-speed. Further, the radius of action of air-meter vanes is usually much less than for cups, so that much higher rotational speeds are involved with the vane air-meter than with the cup anemometer. Thus, whilst the cup anemometer may be used in a wide range of wind-speeds, the vane air-meter as usually designed is not suitable for winds greater than about 15 to 20 m.p.h.

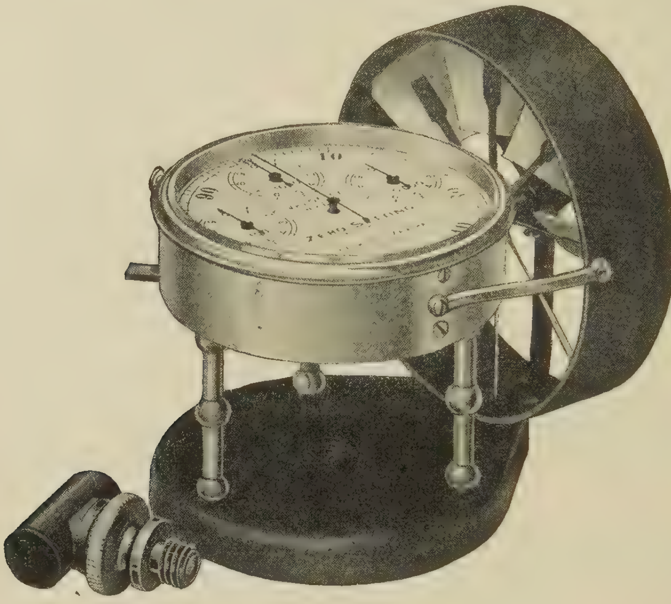


Figure 2. Vane air-meter.

Ower (1926 b and 1937) has given a quantitative treatment of the vane air-meter when the effects of friction are included, making use of the aerodynamic characteristics of flat plates. The analysis shows the effects of the following variables on the motion:—

(i) *Air density.* Wind-speed measurements may be made to an accuracy of 1 per cent by neglecting air-density variations up to about 10 per cent from the standard (calibration) value. Correction for air density is not therefore normally required except at high-level stations. The correction is proportionately greater the lower the wind-speed.

(ii) *Blade angle.* The optimum blade angle—that is, the angle between blade and wind direction for which the anemometer has the best response in low winds—appears to be about 40°.

(iii) *Wind gustiness.* The anemometer over-estimates the mean wind-speed when the wind is variable. For air meters of normal dimensions (diameter of vane circle, 3 in. to 4 in.) the over-estimation is less than 1 per cent for a 10-per-cent simple harmonic variation of speed on

either side of the mean. The error, however, increases with the square of the amplitude of the fluctuations and may, perhaps, be considerable in natural winds.

When setting up an air meter for wind-speed determination it is not necessary to know the wind direction with great accuracy. Tests have shown that with an instrument of normal proportions the error on the indicated speed reached 2 per cent only for 20° yaw.

The vane air meter is an exceedingly useful instrument in the range of low wind-speeds, being relatively compact and simple to use. Like the cup anemometer, it is an integrator of air flow, the mean speed over any time interval being obtained from the difference between initial and final dial readings and the application of the calibration curve for the particular instrument. Although the dial is normally engraved in " feet of air ", accurate work demands the use of a calibration curve which should be appropriate to the attitude in which the air meter is used. The common advice of makers to add a certain number of feet/min. to indicated air speeds to obtain true speeds should not be followed without verification.

§ 8. PITOT-STATIC TUBE

Pitot (1732) was the first to employ the excess pressure in a tube bent so that one end faced into a flow of water as a measure of the speed of the flow. The indications of such an instrument used in liquid or gas were clarified by means of Bernouilli's theoretical investigations (1738) into the motion of an idealized fluid which did not possess viscosity. Bernouilli's equation relating the hydrostatic pressure p_s at a point on a streamline to the velocity v at the point may be written

$$p_s + \tfrac{1}{2}\rho v^2 = \text{constant} = p_t, \text{ say,} \qquad \ldots\ldots(3)$$

where ρ is the fluid density, assumed constant (fluid incompressible), and changes in level are neglected. The equation is a statement of Newton's second law for fluid in steady motion. The second term on the left of equation (3) has the dimensions of pressure and is usually referred to as the *velocity head*, whilst the sum of the two terms is called the *total head*. Thus along any streamline the total head is constant, and where the velocity is high the hydrostatic pressure (usually abbreviated to static pressure) is low, and *vice versa*. The neglect of compressibility will involve no serious error for air speeds less than about 60 m./sec. (140 m.p.h.), and the neglect of change in level is justified in all practically important cases of air-flow measurement.

From equation (3), the total head may be measured by reducing the air speed on a streamline to zero. This is the action of the Pitot tube; assuming no disturbance to the flow by the tube, and neglecting the effect of viscosity at the wall of the tube, the energy of motion of the fluid impinging on the open end is converted into potential energy, represented by a correspondingly increased pressure

within the tube. Barker (1922) has shown experimentally that the neglect of viscosity is justified except for very small tubes at low speeds, or, in terms of the Reynolds number, when this is less than about 60. For the rest it has been found that an open-ended tube, the form of which may be varied within wide limits, gives a true measure of total head.

In order, however, to determine the air speed the static pressure must also be measured, so that from (3)

$$v = \sqrt{\frac{2(p_t - p_s)}{\rho}}. \qquad \ldots\ldots(4)$$

If a small thin plate be introduced into the air stream with its plane in the direction of motion, each side of the plate will experience the static pressure at that position. To obtain a practicable arrangement with a lead to a manometer, and such that no disturbance to the flow is created at the point of measurement, a more complex arrangement is required. The most general procedure is to combine a so-called

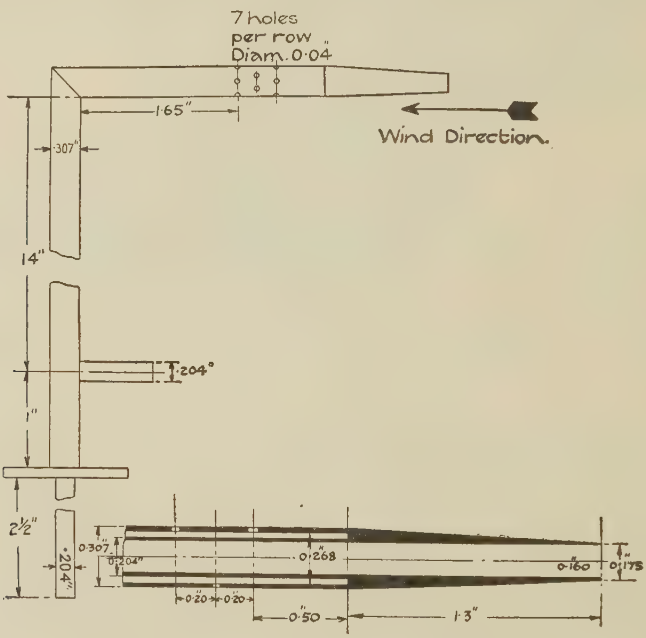


Figure 3. N.P.L. standard Pitot-static tube : from *Modern Developments in Fluid Dynamics*, 1938 (Oxford University Press).

static tube with a total head or Pitot tube. The single instrument, a Pitot-static tube, consists of a double-walled tube bent at right angles, the shorter branch being aligned with its axis in the direction of the flow. This branch is tapered down by a conical or hemispherical fitment at the end to leave a single hole facing the wind, connected to one arm of a manometer through the inner tube. At about six tube diameters behind the total head opening, small holes are drilled through the outer tube with their axes normal to the axis of the tube, and they

connect with the other limb of the manometer through the annular space between the inner and outer tube. The flow being parallel to the walls of the tube at the location of the holes, the latter experience the static pressure there, so that the manometer will indicate the difference between the total-head and static pressure, i.e. the velocity head.

Figure 3 shows the standard N.P.L. design of Pitot-static tube with a conical head. This is rather liable to damage, and another widely used and more robust form, with hemispherical head, is shown in figure 4. Both instruments have been carefully calibrated on a whirling arm in still air and, using equation (4) above, give values of air speed accurate to within 0·5 per cent of v (generally better) over the speed range 3 to 250 ft./sec. It is to be noted that with such

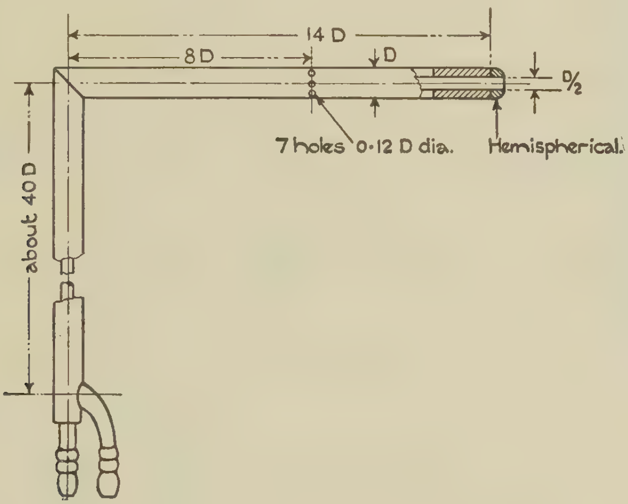


Figure 4. Pitot-static tube with hemispherical head : from *Modern Developments in Fluid Dynamics*, 1938 (Oxford University Press).

instruments the static and total pressures are not measured at the same point, but the gradient of static pressure along the stream will rarely be sufficient to introduce sensible error on this account.

The error in air-speed determination by lack of exact alignment of the Pitot-static tube with the wind direction is less than 1 per cent for an angle of yaw up to 5°, and in the hemispherical nosed instrument is less than $2\frac{1}{2}$ per cent up to an angle of 30°.

The Pitot-static tube is simple in construction, and damage to it likely to involve error in its readings will generally be apparent. If built according to N.P.L. specification it represents a standard instrument against which other types of anemometer may be calibrated. It has the disadvantage that only very small pressure differences are developed for moderate or small air speeds—air of normal density moving at 1 m./sec. (2·24 m.p.h.) gives 0·0062 cm. of water head only. Sensitive yet fairly simple manometers, such as the Chattock-Fry tilting U-tube,

have, however, been devised which are responsive to pressure differences of about 0·0001 cm. of water. For descriptions of such instruments the reader is referred to Ower's *The Measurement of Air Flow.*

If the Pitot-static tube is used in a fluctuating wind whose variations are too rapid to be followed by the manometer, the mean measured pressure difference (which is all that can then be determined) will be in excess of $\frac{1}{2}\rho\bar{v}^2$, where $\bar{v}$ is the mean air speed. For let v' be the instantaneous departure of the air speed from its mean value. The differential pressure at any instant is given by

$$p = \tfrac{1}{2}\rho(\bar{v} + v')^2 \qquad \ldots\ldots(5)$$

and its mean value $\bar{p}$ by

$$\bar{p} = \tfrac{1}{2}\rho(\bar{v}^2 + \overline{v'^2}) = \tfrac{1}{2}\rho v^2\left(1 + \frac{\overline{v'^2}}{\bar{v}^2}\right), \qquad \ldots\ldots(6)$$

since $\bar{v}' = 0$.

If v_i is the indicated mean speed for a pressure difference $\bar{p}$, then

$$\bar{p} = \tfrac{1}{2}\rho v_i^2, \qquad \ldots\ldots(7)$$

so that the ratio of indicated mean speed to the true mean is, from (6) and (7),

$$v_i/\bar{v} = (1 + \overline{v'^2}/\bar{v}^2)^{\frac{1}{2}}. \qquad \ldots\ldots(8)$$

If the fluctuations are approximately simple harmonic, so that we may write $v' = \bar{v}\, a \sin \omega t$ (a = relative amplitude of fluctuations of period $2\pi/\omega$), equation (8) becomes

$$v_i/v = (1 + a^2/2)^{\frac{1}{2}}. \qquad \ldots\ldots(9)$$

When $a = 0{\cdot}2$, v_i is 1 per cent in excess of $\bar{v}$, whilst for $a = 0{\cdot}5$, the error is 6 per cent of $\bar{v}$.

The above treatment requires modification if the speed fluctuations are accompanied by directional variations, as in turbulent flow. In the latter case turbulent components of velocity normal to the mean wind direction will produce impact pressures at the static holes. Theoretical work by Goldstein (1936) and experimental tests by Fage (1936 b) suggest that for isotropic turbulence the factor 2/3 should appear before the term in $\overline{v'^2}$ in (6) above.

§ 9. PRESSURE - TUBE ANEMOMETERS FOR METEOROLOGICAL USE

The total-head tube was applied early, though crudely, to meteorological ends. Dr. Lind, physician to the Court at Windsor, employed in 1775 a U-tube one limb of which was bent over and held "face to wind" by mounting this limb on a vertical spindle so that the U-tube formed a direction vane. Wollaston in a posthumous paper to the Royal Society in 1829 described a modification of Lind's instrument in which the simple U-tube was converted into a 2-liquid differential manometer with a magnification of about 5 times. A more practical instrument was devised by Adie in 1836 and used for several years at the Royal Observatory in Edinburgh. This consisted of a bell-mouthed facing tube communicating with the interior of an inverted cylinder floating in water so that the cylinder rose from its zero position when a wind blew.

In 1892, a year which stands out in the history of experimental meteorology, W. H. Dines described the first form of the pressure-tube anemometer which has taken his name. The instrument is a development of Adie's, but so great was the advance on the latter that in effect it was completely new. It is a self-recording instrument, the modern form of which is shown in figure 5, the head, and figure 6, the recorder. The total-head tube is mounted on a wind vane so as to keep the opening face to wind. Instead of using static holes on the periphery of the total-head tube, as in the Pitot-static tube, a symmetrical array of holes (S, figure 5)

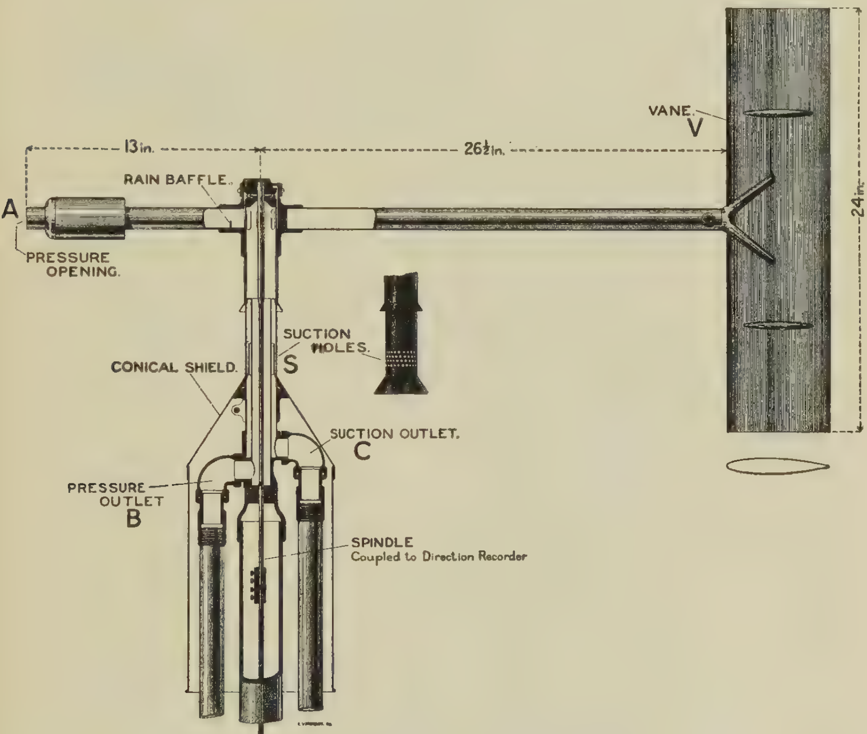


Figure 5. Head of Dines anemometer : from Meteorological Office, *Observers' Handbook*, 1939 (H.M. Stationery Office).

is drilled in the outer wall of a double-walled vertical tube, immediately below the air-tight bearing collar for the head. Air-flow past these holes gives rise in fact to a suction which is independent of wind-direction and augments the recorded pressure difference p by a factor k, so that

$$p = \tfrac{1}{2}(1+k)\rho v^2, \qquad \ldots\ldots(10)$$

where $k = 0{\cdot}49$ for the pattern shown.

The suction holes serve a very important function in this instrument. It might have been supposed that for meteorological purposes static holes or their

equivalent would be unnecessary and that an open limb of the manometer housed out of the wind would experience the appropriate static pressure. This is not the case; small but rapid variations in static pressure are by no means infrequent,

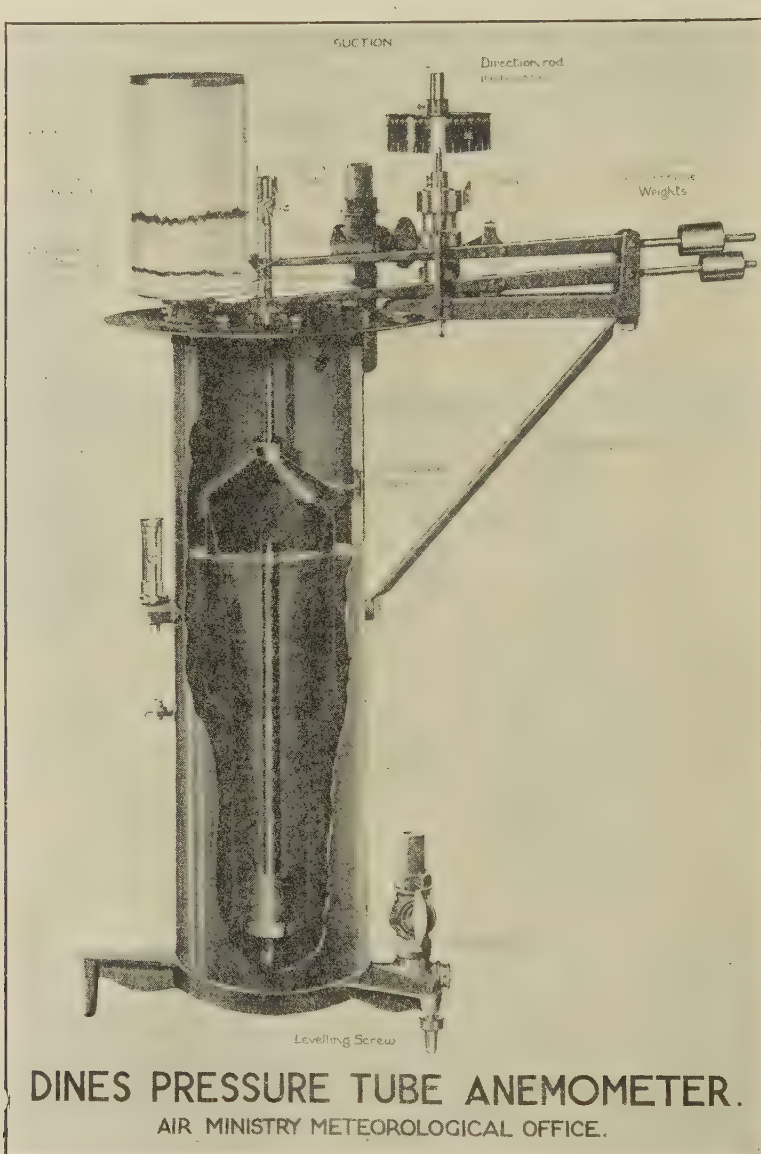


Figure 6. Dines pressure-tube recorder: from Meteorological Office, *Observers' Handbook*, 1939 (H.M. Stationery Office).

as evidenced by a microbarograph, and these variations may be experienced less fully and with lag in a "closed" room relative to the air outside. Further,

(i)

(ii)

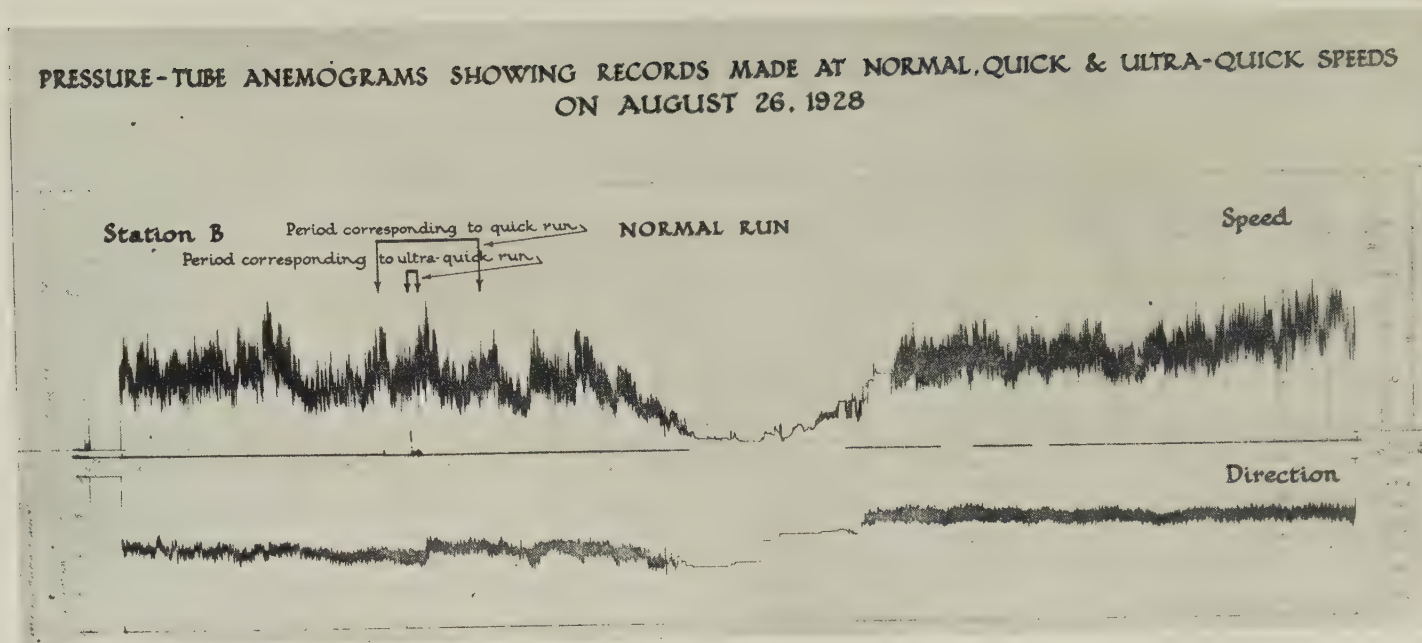


Figure 7. Dines anemographs. This figure (see also overleaf) shows an example of normal, quick and ultra-quick runs made on the same occasion at Cardington, Beds. The first two traces show the speed and direction on the normal time scale (i.e., the record shows 24 hours of wind). On this record two long arrows are marked to indicate the period, 2 hours, during which the quick run was being made and two short arrows to indicate the period, 10 minutes, during which the ultra-quick record was being made. The two upper traces overleaf show the quick-run records for speed and direction (time scale 12 times the normal). The two lower traces overleaf show the ultra-quick record for speed and direction (time scale 144 times the normal).

From Meteorological Office, *Geophys. Mem.* no. 54 (H.M. Stationery Office).

(Traces (iii) to (vi) overleaf.)

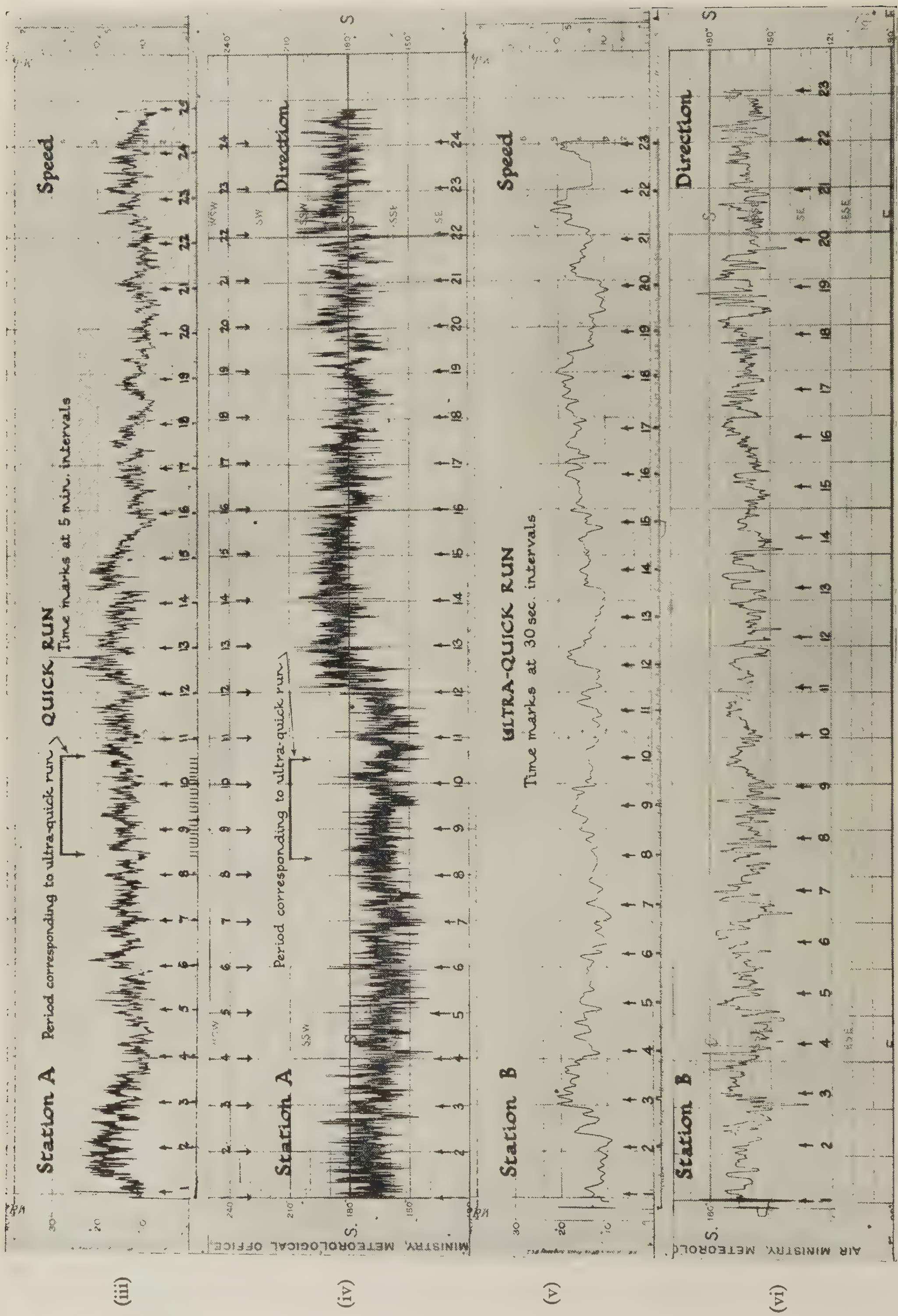


Figure 7 (cont.).

opening and closing of doors, etc., causes appreciable temporary changes in static pressure inside a room. The presence of the suction holes on the head of the instrument ensures that both sides of the manometer will be equally affected by pressure changes in the atmosphere.

The anemometer recorder consists on the wind-speed side of a metal float, cylindrical on the outside, but shaped to a quartic hyperboloid of revolution on the inside. 1-in. metal piping connects the total head and suction holes to the inside of the float and to the float chamber respectively, a maximum of 100 feet separation between head and recorder being possible. The special shape of the float results in the rise of the latter being directly proportional to wind-speed (more precisely proportional to the product $v\rho^{\frac{1}{2}}$) and the scale of the instrument is 1·52 mm./m.p.h. for a standard air density of 0·00125 gm./cm^3. Recording is by means of a pen attached to a rod rising from the float through a nearly air-tight collar on the float chamber.

The direction record is obtained by linking the head through a vertical spindle of metal tube to a double helix mounted on the float chamber and free to rotate in sympathy with the head. Two pen arms are balanced so that one or other, or both, ride up and down inclined planes on the helix as the latter rotates. As the wind vane moves through the north point one pen goes out of action and the other comes into action.

The Dines pressure-tube anemometer thus gives a continuous record of wind-speed and direction, showing the characteristic sequence of gusts and lulls and the continuous variation in direction of which the wind is normally composed except when the air is stably stratified. The one real difficulty in operating the instrument is to obtain good pen-marking when the wind is high and very gusty or the atmosphere very dry. Some typical records, with normal time scale and quick run, are shown in figure 7. A very full account of the pressure-tube anemometer with a discussion of some records has been given by Gold (1936).

No single instrument can be designed to record every quality of the wind, and the pressure-tube anemometer is no exception. It is well to keep in mind the limitations of any instrument, and in the present case they are, briefly, as follows:—

(i) The instrument is not suited to the measurement of very light winds, its indications lacking accuracy below about 3 to 5 m.p.h. At a speed of 1 m.p.h. the vertical force on the float is only $\frac{1}{3}$ gm. weight, whilst at 100 m.p.h., the upper limit of speed allowed for on the standard chart, the corresponding force is $3\frac{1}{2}$ kg. weight.

(ii) Accurate values of the mean wind-speed over a period are difficult to obtain from the record, especially when the gustiness is large. Great accuracy is, however, only seldom required—in the main for special investigations.

(iii) The full amplitude of the more rapid gusts and lulls may not be recorded if the head is very distant from the recorder. With 100 ft. of 1-in.

pipe between head and recorder there is much reduction in recorded amplitude in winds less than 10 m.p.h. if the period of fluctuation is less than 10 sec. (but see (iv) below). At 60 m.p.h. a faithful record is obtained for fluctuations of period greater than 2 sec. (Giblett *et al.*, 1932).

(iv) Resonance effects may be shown on the speed and direction traces for certain periodicities in speed and direction variation. In each case the period for resonance, of the order of 1 sec., increases with decrease in wind-speed.

Graham (1936) has successfully employed a differential aneroid recorder to obtain a more complete record of the gusts and lulls than would be given by the float recorder. So little air motion is involved in the movements of the aneroid diaphragm that the damping action of the pressure leads is no longer in evidence. The aneroid recorder may be useful at sea, where the float recorder is, of course, useless.

§ 10. VENTURI TUBE, PLATE ORIFICE AND SHAPED NOZZLE

There is a further class of instrument deriving from Bernouilli's equation which finds frequent application to the measurement of air flow through pipes. If the velocity of flow through a pipe is increased locally by a constriction in the pipe, the increase in the velocity must be associated with a pressure gradient down the pipe from the section where the velocity is low to where it is high. Bernouilli's equation gives

$$p_1+\tfrac{1}{2}\rho v_1^2=p_2+\tfrac{1}{2}\rho v_2^2, \qquad \ldots\ldots(11)$$

where the subscripts refer to the full bore and the constriction respectively and ρ is assumed constant. Further, the ratio v_1/v_2 is equal to the ratio a_2/a_1 of the cross-sectional areas at the constriction and at full bore, so that equation (11) gives

$$v_2=\left(\frac{2(p_1-p_2)}{\rho(1-a_2^2/a_1^2)}\right)^{\frac{1}{2}}. \qquad \ldots\ldots(12)$$

In fact the neglect of viscosity in this treatment gives a greater value of the pressure difference than is realized, which leads to the introduction of a discharge coefficient α, less than unity, on the right-hand side of (12), whose value has to be determined experimentally for the particular design of constriction employed. Again, the neglect of compressibility is only justified if p_1/p_2 is very near unity ($p_1/p_2<1{\cdot}02$), but its effect can be taken into account at the cost of a more cumbersome expression than (12).*

The merit of a constriction resides in the much greater pressure difference

* The more general form of Bernouilli's equation is $v^2/2+\int\frac{dp}{\rho}=$const., giving $\frac{v_2^2-v_1^2}{2}=\int_{p_2}^{p_1}\frac{dp}{\rho}$, and the variation in ρ, assumed to take place adiabatically, is allowed for by substitution from $p/\rho^\gamma=$const.

which may be produced by this means than by the Pitot-static tube, so that a simple U-tube may be used in the measurement of quite small air speeds. The ratio of the pressure differences in the two cases is equal to $\alpha^2(a_1{}^2/a_2{}^2-1)$, which, for a constriction half the full diameter of the pipe and a discharge coefficient near unity, is equal to about 15.

The velocity of air in a pipe varies from zero at the wall to a maximum on the pipe axis, the form of the variation at sections sufficiently far from the tube mouth or from bends depending on the Reynolds number of flow and the wall roughness (Goldstein, 1938). For this reason the total flow through a pipe can be obtained from Pitot-static tube readings only after an exploration has been made of the velocity profile across a tube section. The constriction method of measurement gives, however, the mean velocity across a tube section—that is its purpose—and it is in terms of the mean velocity that the discharge coefficient is determined.

Various designs of constriction are employed. The earliest, due to Venturi at the end of the 18th century, consisted essentially of two truncated cones, narrow ends together. In the modern design of the Venturi tube the angle of convergence of the upstream cone is usually about 20°, though this figure is not of

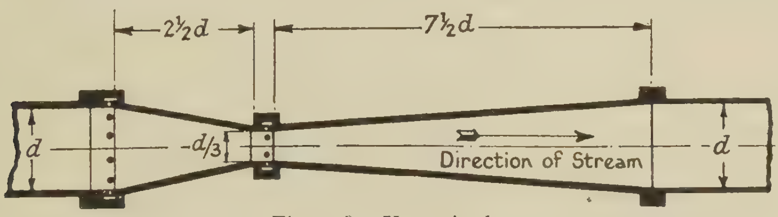


Figure 8. Venturi tube.
From Ower, *Measurement of Air Flow*, 1933 (Chapman & Hall).

great importance. The angle of divergence of the downstream member must, however, be small, 6° being a suitable figure, if a net loss in pressure head is to be avoided. The ratio of throat diameter to full-bore diameter is usually between $\frac{1}{2}$ and $\frac{1}{3}$, and a design, as shown in figure 8, conforming to the above figures, has a discharge coefficient between 0·95 and 0·98, the actual value increasing with the Reynolds number. The pressure difference is taken between the upstream end of the convergent cone and the throat, from small static holes bored in the periphery of the tube at the two sections.

In recent years two other designs of constriction, the plate orifice and shaped nozzle, have largely superseded the Venturi tube, being more compact, less expensive and less troublesome to insert into an existing pipe-line than the latter. They are shown in figures 9 and 10, which are mainly self-explanatory. The plate orifice is the simpler, but has a discharge coefficient of about 0·60 only, the actual value of α as a function of Reynolds number for the pattern shown in figure 9 being obtainable from tables (see Ower, 1933). A further disadvantage is that a considerable fraction of the pressure drop at the constriction is not recoverable, so that greater power must be expended to maintain a given flow

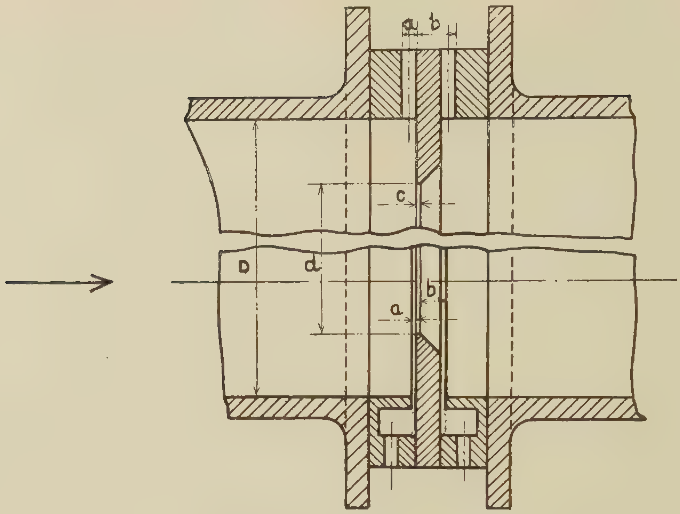


Figure 9. German standard plate orifice, 1930.
From Ower, *Measurement of Air Flow*, 1933 (Chapman & Hall).

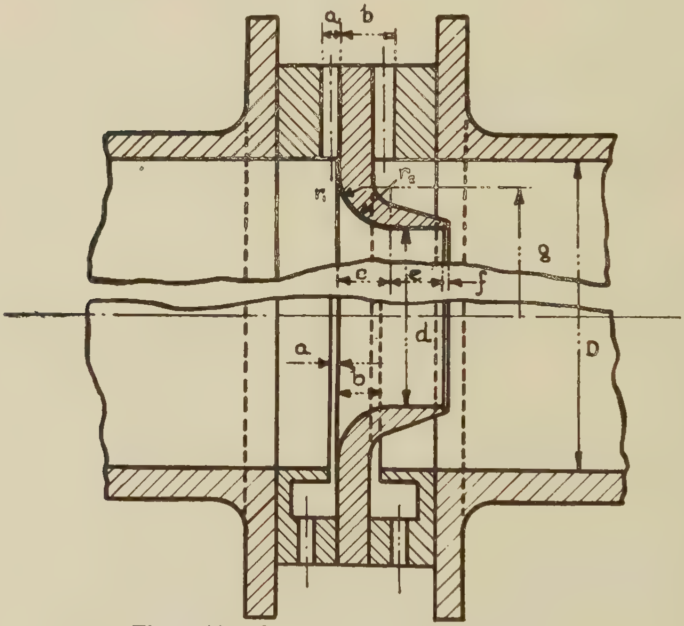


Figure 10. German standard nozzle, 1930.
From Ower, *Measurement of Air Flow*, 1933 (Chapman & Hall).

The pressures are tapped off either through two single pressure holes as shown on upper portion of figures 9 and 10 or, as shown below, from two annular chambers in communication with the pipe pressure over the whole circumference.

after the insertion of an orifice plate. The shaped nozzle has a much higher discharge coefficient, 0·95 to 0·98, and for a given pressure drop shows a smaller resultant loss of head than the plate orifice.

A marked advantage attaching to the use of the plate orifice or shaped nozzle is that for a standard design suitably placed in the pipe no special calibration is necessary. For full details of their characteristics the reader is referred to Ower's valuable text-book.

§ 11. PRESSURE-PLATE ANEMOMETERS

In this class of instrument, the wind-pressure exerted on a plate, generally rectangular, is balanced against a restoring force induced by the deflection of the plate, the restoring force arising from gravity, spring compression or simple torsion. The gravity-controlled type of instrument—the swinging plate or pendulum anemometer—is the first anemometer of which there is any record. A certain Egnatio Danti erected swinging plates for the measurement of natural winds at Bologna and Florence about 1570, whilst Hooke is associated with its inception in this country some time prior to 1667. It is not now in general use, but very light plates having a good response to fluctuating winds have been used by Schmidt (1929) and Scrase (1931) for investigations on atmospheric turbulence, cinematograph recording being employed.

Osler (1836) devised, and later (1862) improved on, a spring-controlled recording pressure-plate held normal to the wind, which was in wide use in this country during last century. Of more interest to modern workers is a well-designed recording pressure-plate with a streamlined housing due to Sherlock and Stout (1931 and 1937). In this the plate movement is very small and is controlled so as to avoid errors due to inertia. Recording is by means of an oscillograph through which passes an out-of-balance current induced in a magnetic bridge by movement of the plate. A number of such instruments have been used in a spatial array to give a picture of the pattern of turbulent atmospheric flow in high winds.

In all the above instruments except the last, uncertain inertial effects have made the interpretation of the records of wind rather uncertain.

Of the torsion type, mention should be made of a very sensitive instrument due to Rees (1927), suitable for use in ducts, etc., for wind-speeds from 10 to 180 ft./min. (5 to 100 cm./sec.). It is a null instrument in which an asymmetrically mounted plate is maintained in a vertical plane by applying torsion to the mounting wire. In consequence it is suited to the measurement of steady winds only.

The force exerted on a plate by the wind is not accurately proportional to the area of the plate nor is it independent of the aspect ratio. It is possible that the degree of turbulence of the approaching air stream also affects the force. On these accounts the characteristics of a pressure-plate system can only be deduced after a wide appeal to the literature of experimental aerodynamics, and calibration is desirable in any case.

§ 12. HOT-WIRE ANEMOMETERS

The class of instrument which depends on the dissipation of heat from a body placed in an air stream, in particular the hot-wire anemometer, has acquired increasing importance in recent years, especially in investigations into the details of turbulent flow. The fundamental research on the dissipation of heat from cylindrical wires is due to King (1914), who provided a theoretical treatment of the problem for cylinders of small diameter-to-length ratio, the results of which were confirmed by experiment.

If a wire is held normal to an air stream of speed v and temperature T_0, and is electrically heated by a current i to bring the wire to a resistance R and temperature T, King's equation gives

$$Ri^2 = (K + c\sqrt{v})(T - T_0), \qquad \ldots\ldots(13)$$

where K is a constant determined by the radiative and convective heat loss from the wire in " still " air and c a second constant depending on the physical properties of the air * and the diameter of the wire. For anemometric purposes the hot wire is generally used in one of two ways, at constant resistance (constant temperatures excess for T_0 constant) or else at constant current.

For the constant-resistance method, equation (13) may be written

$$i^2 = i_0^2 + a\sqrt{v}, \qquad \ldots\ldots(14)$$

where i_0 is the value of i when the wire is in still air and a is a constant. From this relation it is seen that the current varies relatively rapidly with air speed for low values of the latter and slowly for the higher speeds. Indeed, the hot-wire anemometer can be made the most sensitive of all types for measurement of low wind-speeds. The method involves the use of a suitable resistance bridge and current-measuring device, and calibration of a particular wire involves simply the determination of i_0 and a.

The constant-current method is simpler experimentally than the constant-resistance method. A sufficiently large resistance is included in the circuit to prevent the resistance variation of the hot wire itself being of any effect, and the potential drop across the wire is used as the index of air speed. The method is not, however, as accurate in general as the constant-resistance method.

A hot-wire anemometer usually consists of platinum wire, occasionally electrolytically pure nickel, stretched between tapered prongs mounted on a base suitable for fixing the anemometer as required. The length of the hot wire can be suited to the particular use to which the instrument is to be put, but if there is appreciable variation in air speed over the length of the wire, some error in mean speed over this distance may be incurred from a calibration made in a uniform flow. Lengths of wire varying from about 1 mm. to several cm. have been used. Again, the wire diameter chosen will depend on the particular application (see below). but will usually be between about 0·1 mm. and 0·002 mm.,

* The constant c involves $\sqrt{\rho}$, so that a hot-wire anemometer gives a measure of the product ρv.

the finer wire being prepared by the Wollaston process. Annealing for several hours at red heat is necessary.

Liability to fracture and to change in calibration, particularly the latter, are the bugbears of hot-wire anemometry. The calibration inconstancy is greater the thinner the wire (greater ratio of surface to volume of wire) and the more elevated the temperature at which it is used. It is advisable to keep the latter below 500° C. The reasons for the variation in heat transfer from the surface of the wire to the surrounding medium under apparently constant conditions are not yet fully understood. Accretion of dust on the wire is considered by some to be important, though the dust-free space effect round a hot body would seem to oppose this interpretation. Possibly surface gas-layer effects are of importance.

Probably the greatest merit of the hot-wire instrument is its ability to respond to rapid changes in wind-speed when fine wire is used. This feature makes the instrument particularly suited to investigations of turbulent flow when used in conjunction with an appropriate recording device—oscillograph, string galvanometer, etc. Dryden and Keuthe (1929) have shown that the ability of a wire to follow variations in air speed depends on a quantity M, the time constant of the wire, given by

$$M = 4{\cdot}2d\frac{A^2 s(T-T_0)^2}{\sigma_0 i^2}, \qquad \ldots\ldots(15)$$

in which

d = density of wire,

A = cross-section of wire,

s = specific heat of wire,

σ_0 = specific resistance of wire at temperature T_0.

Thus, if the air speed is changed suddenly from a steady value at which the wire resistance is R_1 to a second steady value at which the corresponding resistance is R_2, the temporal change in resistance R is given by

$$R - R_2 = (R_1 - R_2)e^{-t/M}, \qquad \ldots\ldots(16)$$

and the smaller the value of M the more rapid the response of the wire. From (15), for a given value of $(T-T_0)$, M decreases with increase in wind-speed, since i is increased to maintain the temperature excess, and increases as the cross-section of the wire—not as A^2, since i must be varied as $\sqrt{A}$ for a given $(T-T_0)$. Again, for a given wire, M is proportional to $[1+\alpha(T-T_0)]$, where α is the temperature coefficient of resistance of the wire, as can be shown by substituting in (15) for i^2 from (13). Thus for rapid response it is advantageous to use as fine wire and as low a temperature excess as the sensitivity of the apparatus will allow.

If the fluctuation in wind-speed be simple harmonic of period $2\pi/\omega$, and if the amplitude be small enough to allow of the resistance variation being assumed proportional to the speed variation, the recorded amplitude of speed fluctuation will be $(1+\omega^2M^2)^{-\frac{1}{2}}$ of the true and will be retarded in phase by $\tan^{-1}\omega/M$. A platinum wire 0·0025 mm. (0·0001 in.) diameter may be used so as to respond with reasonable accuracy to fluctuations of about 100 per sec. This is probably

near the upper limit of the frequency of natural wind fluctuations, but is much less than the majority of frequencies present in wind-tunnel flow. Automatic compensation for the response at the higher frequencies may be obtained by the use of a suitable valve circuit (Dryden and Keuthe, 1929 ; Simmons, 1934) in which the amplification of applied voltage fluctuations increases with increase in frequency of the latter.

Various methods have been used to compensate a hot-wire anemometer circuit against variations in air temperature (Davis, 1921), whilst the wide departure from linearity in response of the hot wire with variation in wind-speed may be overcome to a considerable extent by the inclusion of a volt-thermometer in the hot-wire circuit (Huguenard *et al.*, 1923).

The sensitivity of a hot wire falls off as the wire is rotated out of the normal position into a line parallel to the flow, the response in the latter position depending in part on the magnitude of the normal components of turbulent velocity. It would appear that the directional properties of the hot wire make it suited to the design of a recorder of the normal components of turbulent velocity as well as of the longitudinal component ; no entirely satisfactory apparatus of the kind is, however, known to the writer.

To summarize, the hot-wire anemometer may be profitably used for the measurement of low wind-speeds, of wind-speed fluctuations, for which purpose it is pre-eminent, and at points near a boundary surface where other instruments would be unsuitably large. Otherwise its fragility and inconstancy of calibration make it an instrument to be avoided.

§ 13. KATA THERMOMETER

The cooling action of a wind on a thermometer raised above air temperature was first used and described by Leslie (1804). Little application of the method was made until L. Hill (1912 and 1922) devised the so-called kata-thermometer for measuring ventilation. The instrument consists simply of a thermometer having a large cylindrical bulb (length 4 cm., diameter 2 cm.) and a thick-walled capillary tube on which the levels corresponding to 38° C. and 35° C. are clearly marked. The thermometer is initially warmed to about 50° C., set up in the position in which the air speed is required to be known, and the time t taken for the mercury to fall from the upper to the lower marked temperature. The air speed v may then be found from the equation (a nomogram will normally be used)

$$B^2 v = \left(\frac{F}{(36{\cdot}5 - T_0) - A}\right)^2, \qquad \ldots\ldots(17)$$

where F is the heat loss from the thermometer in cooling from 38° C. to 35° C., T_0 the air temperature and A and B constants to be determined experimentally.

The instrument is suited to the measurement of low wind-speeds, below about 1 m./sec. Its limitation is obvious—it gives only the mean speed for the period in which the given temperature drop occurs.

§ 14. THE MEASUREMENT OF WIND DIRECTION

The preceding sections have been mainly concerned with the measurement of air speed, occasional reference only having been made to the determination of direction. The need for instrumental measurement of the latter arises first in meteorology, second in the determination of local deviations from a general direction of flow brought about by a solid body placed in an air stream, and also in investigations of the variations in direction associated with turbulent flow.

The need for an instrument which should indicate the horizontal direction of the wind must early have been felt by the mariner, and so it is no surprise to find that the wind vane is of great antiquity. Ornate vanes were used and special buildings were erected in the Greek and Roman world to show the wind direction, the most memorable of the buildings being the Tower of the Winds at Athens with its eight faces directed to Boreas, Zephyr and their kind. The registering wind vane appears to have been introduced in the seventeenth century by Wren, who linked a recording clockwork mechanism to the weather cock, the Christian symbol of medieval times which had largely replaced the triton of classic days. In more recent times the wind vane for scientific usage has been a flat plate, splayed plate, or today, a symmetric aerofoil (see figure 5) or pair of asymmetric aerofoils with similarly cambered sides facing.

A recording vane should be designed so that the restoring force on the vane when it is out of wind is a maximum, the inertia of the vane and recording mechanism should be a minimum consistent with strength and durability and it should be well damped yet responsive. Resonance of the vane for the direction periodicities involved is to be avoided. In fact, the natural period of a vane decreases with increase in wind-speed, so that the vane response varies with the latter. The natural damping of most vanes is effectively independent of wind-speed. Giblett and others (1932) have treated the characteristics of the symmetric aerofoil vane and recorder used in the modern Dines anemograph, whilst Grunow (1935) has designed a compound asymmetric aerofoil vane which has excellent qualities. A wind vane should be well balanced, for lack of balance in a vane mounted on a spindle which is not truly vertical will lead to incorrect indications in a light wind.

The measurement of a local steady wind direction in the region of a body placed in an air stream may be roughly determined by the use of wool tufts, smoke, etc. For accurate measurements an instrument due to Lavender (1923) makes application of the fact that the rate of variation of head on a total-head tube with angle of yaw is a maximum at about 45°. Two total-head tubes are mounted with their axes at right angles in a horizontal plane, the tube mouths being towards but short of the intersection of the axes. The assembly is rotated about a vertical axis until there is no difference of pressure at the tube mouths and the setting of the horizontal axis of symmetry gives the horizontal direction of the wind. A second similar pair of tubes in the vertical plane is used to obtain the inclination of the wind to the horizontal. Accuracy of direction to 0°·1 may be obtained by this means.

The hot-wire direction meter of Simmons and Bailey (1926) is also a null instrument in which the heat loss from two identical wires set at an angle of 10° to each other is equal when the bisector of the angle between them lies in the wind direction. Equality of heat loss is indicated when the bridge in which the wires form opposite arms is in balance.

§ 15. WIND DIRECTION AND WIND STRUCTURE

The instantaneous wind direction at a point in a field of turbulent flow whose mean speed is $\bar{u}$ is related to the turbulent components of velocity as follows. Take rectangular co-ordinate axes with the x-axis in the direction of mean flow. Let the instantaneous wind vector V, represented by the line OP in figure 11, have components $u(=\bar{u}+u')$, v', w' in the x, y, z co-ordinate directions. The direction of V relative to the mean direction is specified by the azimuth θ and inclination ϕ (figure 11). Then

$$\left.\begin{aligned} u' &= V\cos\theta\cos\phi, \\ v' &= V\sin\theta\cos\phi, \\ w' &= V\sin\phi, \end{aligned}\right\} \qquad \ldots\ldots(18)$$

or

$$\left.\begin{aligned} v' &= u\tan\theta, \\ w' &= u\sec\theta\tan\phi, \end{aligned}\right\} \qquad \ldots\ldots(19)$$

Thus the simultaneous measurement of V, θ and ϕ determines u, v', w' and from a series of observations $\bar{u}$ and u' may also be deduced. Alternatively u, θ and ϕ may be measured and u', v', w' deduced.

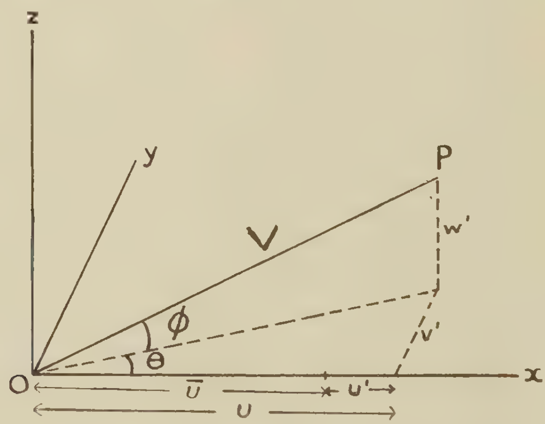


Figure 11.

Difficulty is experienced when attempts are made to measure simultaneously all the relevant quantities at a single point or small domain, and neither for the wind tunnel, pipe, etc., nor for the atmosphere has adequate apparatus yet been developed. Some progress has, however, been made. Thus the ultramicroscope in the hands of Fage (see § 3) has been successfully applied to the measurement of the maximum numerical values of u', v', w' for water-flow through a tube.

Universal vanes have been developed by Taylor (1927), Becker (1930), Scrase (1931), and Best (1934) for recording the variation in wind direction (azimuth and elevation) in the atmosphere. In order to interpret their records, Scrase and Best made simultaneous measurements of mean wind-speed at a nearby point, using equations (19) in an approximate form ($\bar{u}$ for u) to give values of v' and w'. The error involved in this procedure may be appreciable since u' may be a considerable fraction of $\bar{u}$ in the lower atmosphere. Further, there is frequently a correlation between the turbulent velocity components, so that, for example, positive values of u' (horizontal component in direction of mean wind) are normally associated with negative values of w' (vertical component), and this correlation will lead to error in deriving arithmetic mean values from equation (19) when $\bar{u}$ replaces u.

Kampé de Férieth has made an interesting application of pressure-tube principles for the evaluation of the time variation in total velocity and direction of flow in the atmosphere. The instrument has been described by Walker (1937 and 1938) and consists of a sphere, 8 cm. in diameter, with a Venturi tube running horizontally through the middle of it and small static holes distributed uniformly over its surface. The sphere is set up with the Venturi axis aligned with the mean wind. The latter cannot, in fact, be known accurately beforehand, but this is of no consequence for the particular design of instrument. Three pressure differences are recorded from suitable combinations of holes as follows :—

(i) The pressure difference p_1 between the throat of the Venturi and the mean pressure at the surface holes, from which

$$p_1 = 1{\cdot}37q, \qquad \ldots\ldots(22)$$

where q is the velocity head, $\frac{1}{2}\rho u^2$. Thus u is determined (figure 11).

(ii) The pressure difference p_2 between two surface holes symmetrical in relation to the Venturi tube in the horizontal plane, from which

$$p_2 = 0{\cdot}05q\theta, \qquad \ldots\ldots(23)$$

where θ is the wind azimuth (figure 11).

(iii) A pressure difference, p_3, similar to p_2, from surface holes in the vertical plane, from which

$$p_3 = 0{\cdot}05q\phi, \qquad \ldots\ldots(24)$$

where ϕ is the inclination of the wind (figure 11).

No error greater than 3 per cent. in either θ or ϕ is involved if the wind direction does not deviate by more than 45° from the Venturi axis. The pressure recorder, involving a diaphragm manometer, gives resolution of fluctuations down to 0·5 sec. periodicity.

Such an apparatus is too large for wind-tunnel investigations; in the wind tunnel the scale of turbulence is too small for any practicable size of sphere, and the frequency of the fluctuations too high to be recorded by a diaphragm manometer.

For many purposes it is not necessary to know the value of the turbulent components at every instant of time. More frequently the root-mean-square value of each component is required or the mean values of their products, e.g. $\overline{u'w'}$, $\overline{u'v'}$. To obtain these quantities from a time record of the fluctuations is very laborious, and electrical methods have been adopted to measure some of them directly. The hot-wire anemometer is specially suited to this type of measurement. A hot-wire circuit and associated amplifier may, for example, be made to give an output proportional to u', and this may be passed through an A.C. milliammeter to give the root-mean-square value of the fluctuation, or again, outputs from two circuits may be passed through an electrodynamometer to give the mean value of a product of two fluctuating entities (Simmons, 1938). It is probably along these lines that a more complete knowledge of the various quantities characterizing turbulent flow will be obtained.

§ 16. ANEMOMETER CALIBRATION

Reference has frequently been made in the preceding sections to anemometer calibration. Here some general observations on the matter will be added.

It cannot be too much stressed that the accurate determination of wind-speed depends in every case, except when the time-displacement method is used, on an appropriate calibration of the instrument concerned. This is because the theory of turbulent motion is at present relatively undeveloped in spite of much research and some significant advance in recent years. Thus the results of theoretical treatment can at present be applied only as a lead in instrument design; for complete knowledge of an instrument's performance, experimental calibration remains necessary. In some few cases standardized patterns of instrument have been adopted and thorough calibrations made. In such cases, e.g. the Pitot-static tube, Dines pressure-tube anemometer, plate orifice, etc., calibration of individual instruments is not necessary. Some caution must be exercised, however, even in these cases. The conditions of calibration may not have been the same as those in which an instrument is to be used (see below). Failure to appreciate such differences has undoubtedly led to error in many cases, and when the differences have been appreciated, lack of a suitable experimental technique for the conditions of use has often hindered adequate calibration.

Two methods of anemometer calibration are effectively available:

(i) by moving the instrument at controlled speed through still air;

(ii) by fixing the instrument in an air stream whose speed is controllable and known.

The second method derives from the first since the speed of the air stream has to be deduced from an instrument calibrated according to the first method—the absolute time-displacement method of measuring air flow is not normally practicable or susceptible of great accuracy.

In the first method, linear translation of the instrument is not usually possible, and a rotating (whirling) arm is used instead. The arm sets up a swirl of air

in its path for which allowance may be made with sufficient accuracy. Unfortunately for the application of the calibration thus obtained, the response of an instrument at given speed on the whirling arm is not necessarily the same as for air moving past a fixed instrument with the same relative speed. In the latter case the air flow may be, and frequently is, turbulent, and this may affect the indications of an instrument in a number of ways. The effect of turbulence on the Pitot-static tube has been considered already (§ 8). This is the most important case, since in practice the second of the above methods of calibration is usually employed, the air speed being determined by means of a Pitot-static tube originally calibrated on the whirling arm.

The second method makes use of a wind tunnel of diameter at the working section sufficient to prevent disturbance to the flow from the presence of the instrument to be calibrated. It is to be noted that a wind tunnel is designed to give a wide core of air in which the motion is uniform—the velocity gradient is confined to a region very near the walls. The Pitot-static tube may be used alongside the anemometer for test, or the air speed may be deduced from the static pressure at the working section, a previous calibration having been made of the pressure at a static hole in the wall of the tunnel in terms of the corresponding Pitot-static tube pressure.

The degree of turbulence in the air flow in tunnels varies a good deal, but in a well-designed tunnel will be sufficiently small to make its effect on the Pitot-static tube of no practical importance. Thus the instrumental calibration will be valid for use in an air stream of the same turbulence as that of the tunnel. Very often, however, instruments will be used in air streams whose characteristics are markedly different from those of the calibration tunnel. This change may effect appreciably the motion of a cup anemometer (see, for example, Spilhaus, 1934). Large-scale turbulence which characterizes the atmosphere gives rise to different effects which for convenience may be referred to as the effects of a gusty wind. Unfortunately it is very difficult to produce a wind of controlled gustiness, so that for knowledge of its effects reliance must to a large extent be placed on theory. For almost all instruments of practical importance, the indicated mean speed of a gusty wind is in excess of the true mean value.

§ 17. ANEMOMETER EXPOSURE

If a body is placed in an air stream, the general flow will be disturbed, in many cases relatively little upwind and to the side of the body, but appreciably downwind. Thus, when using an anemometer for obtaining the free flow over a surface, care must be taken in positioning it and the observer who is operating the instrument. If obstacles are present whose detailed effects are not required to be known, the anemometer should be placed well upstream or even further downstream of them to avoid their wakes. The observer will refrain from standing any nearer to the instrument than necessary, certainly not upwind of it, and preferably as far downwind or to the side as possible.

Reference has been made in § 2 to the variation of wind with distance from a boundary. Not only is the variation with distance less at increasing distance, but the effects of surface obstacles upwind of the anemometer location will generally be less at the greater distance from the boundary. It is well, therefore, to expose the anemometer as far from the surface as possible in order to obtain a reading representative of its distance.

§ 18. SUMMARY OF INSTRUMENTS SUITED TO DIFFERENT KINDS OF AIR-FLOW MEASUREMENT

It may be of use to conclude this article with a list of those instruments most suited to different kinds of air-flow measurement. The choice is to some extent personal, and no claim is made for its being exhaustive. It is based on an experience of every instrument listed. Where the word " speed " is used, a temporal mean value is to be understood unless otherwise stated.

Nature of measurement	*Favoured instruments*
Pipe flow	
Local speed, velocity profile.	Pitot-static tube of appropriate dimensions. Hot-wire anemometer.
Mean speed over pipe section.	Plate orifice or shaped nozzle. Side static hole (calibrated against Pitot-static tube traverse) where sufficient head available.
Wind tunnel	
Speed of undisturbed flow in uniform " core ".	Pitot-static tube. Side static hole (frequently calibrated against Pitot-static tube) where sufficient head available. Vane air-meter or small cup anemometer may be useful for low speeds.
Speed in neighbourhood of solid body placed in flow.	Small Pitot-static tube or hot-wire anemometer.
Speed fluctuations. Many special investigations will need apparatus specially adapted.	Hot wire anemometer (special circuit).
Ducts, etc.	
Local speed.	Vane air meter.
Atmosphere	
Record in time of instantaneous speed and direction.	Dines pressure-tube anemograph.
Speed over limited times. Velocity profile near ground.	Vane air meter for low or moderate wind-speed. Portable cup anemometer for all speeds normally encountered.
Speed fluctuations of short period.	Recording hot-wire anemometer.
Ventilation in factories, etc.	Kata thermometer.

ACKNOWLEDGEMENTS

I wish to express my indebtedness and thanks to the following for permission to reproduce illustrations to the text:—

Messrs. Negretti and Zambra for figure 2; the Oxford University Press for figures 3 and 4; the Director, Meteorological Office, London, and the Controller of H.M. Stationery Office for figures 5 to 7; Messrs. Chapman and Hall for figures 8 to 10.

REFERENCES

General :

The Measurement of Air Flow, by E. Ower (Chapman & Hall, London), 2nd edition, 1933, treats the basic principles of air-flow measurement and gives a full description of anemometers used for industrial purposes.

Modern Developments in Fluid Dynamics, in 2 volumes, edited by S. Goldstein (Oxford University Press), 1938, contains a section (VI, vol. 1, pp. 234–94), "Experimental Apparatus and Methods of Measurement", of particular interest to the researcher concerned in the measurement of fluid flow.

Handbuch der Meteorologischen Instrumente, edited by E. Kleinschmidt (Berlin, Julius Springer), 1935, contains a section (VI, pp. 331–96), "Wind Measurement at the Earth's Surface", which gives a very adequate survey of meteorological anemometry.

Experimental Studies of Anemometers, by S. P. Fergusson, Harvard Meteorological Studies No. 4 (Harvard University Press), 1939. This contains a valuable bibliography of 264 papers.

Dictionary of Applied Physics, Vol. III (Macmillan), 1923. Articles on "Meteorological Instruments" and on "Meters".

Historical Surveys of Meteorological anemometry have been given by :

J. K. Laughton, 1882. *Quart. J. R. Met. Soc.* **8**, 161–89.

Cleveland Abbe, 1888. *Treatise on Meteorological Apparatus and Methods* (Washington, U.S.A.), 201–309.

R. Bentley, 1905. *Quart. J. R. Met. Soc.* **31**, 187-9.

Additional items :

Adie, 1836. *New Philos. J. Edinburgh*, **22**, 309.
Barker, 1922. *Proc. Roy. Soc.* A, **101**, 435–45.
Becker, 1930. *Met. Z.* **47**, 183.
Beckley, 1856. *Rep. Brit. Ass.* **2**, 38.
Best, 1935. London, Meteor. Office, *Geophys. Mem.* no. 65.
Brewster, 1830. *Edinburgh Encyclopaedia*, "Anemometers".
Cacciatore, 1840. *Ann. Soc. Met. Ital.* **1**, 201.
Chree, 1895. *Phil. Mag.* **40**, 639–40.
Davis, 1921. *Proc. Phys. Soc.* **33**, 152.
Dines, 1892. *Quart. J. R. Met. Soc.* **18**, 165–85.
Dryden. Reported by Marvin, 1934. *Mon. Weath. Rev.* **62**, 115–20.
Dryden and Kenthe, 1929. *Rep. Nat. Adv. Com. Aeron., Wash.*, no. 320.
Fage, 1936 a. *Phil. Mag.* **21**, 80–105.
Fage, 1936 b. *Proc. Roy. Soc.* A, **155**, 576–96.
Fage and Townend, 1932. *Proc. Roy. Soc.* A, **135**, 656–77.
Giblett *et al.*, 1932. London, Meteor. Office, *Geophys. Mem.* no. 54.
Gold, 1936. *Quart. J. R. Met. Soc.* **62**, 167–206.

GOLDSTEIN, 1936. *Proc. Roy. Soc.* A, **155**, 571–5.
GRAHAM, 1936. *Rep. Memor. Aero. Res. Comm., Lond.*, no. 1704.
GRUNOW, 1935. *Z. InstrumKde*, 55.
HILL, 1912. *Heating and Ventilating Mag.*
HILL *et al.*, 1922. *Proc. Roy. Soc.* B, **93**, 198.
HOOKE, 1667. *Philos. Trans.* **2**, 24 and 444.
HUGUENARD *et al.*, 1923. *C.R. Acad. Sci., Paris*, **176**, 287–9.
JOHNSON and HEYWOOD, 1938. London, Meteor. Office, *Geophys. Mem.* no. 77.
KAMPÉ DE FÉRIET. Reported by WALKER, 1937. *Quart. J. R. Met. Soc.* **63**, 497–8.
KAMPÉ DE FÉRIET. Reported by WALKER, 1938. *Ibid.* **64**, 625–9.
KING, 1914. *Philos. Trans.* **214**, 373.
LAVENDER, 1923. *Rep. Memor. Aero. Res. Comm., Lond.*, no. 844.
LESLIE, 1804. *Experimental Enquiry into the Nature of Heat*, p. 282.
MARIOTTE, 1717. *Œuvres*, **2**, 406.
MAURER, 1883. *Z. InstrumKde*, **3**, 189–92.
OSLER, 1836. *Rep. Brit. Ass.* **2**, 33.
OWER, 1926 a. *J. Sci. Instrum.* **3**, 109.
OWER, 1926 b. *Phil. Mag.* **2**, 881.
OWER, 1937. *Ibid.* **23**, 992-1004.
PATTERSON, 1926. *Trans. Roy. Soc. Can.* **22**, 1053.
PITOT, 1732. *Mém. Acad. Sci., Paris.*
REES, 1927. *J. Sci. Instrum.* **4**, 311.
ROBINSON, 1846. *Rep. Brit. Ass.* 111.
SCHMIDT, 1929. *Met. Z.* **46**, 495.
SCHRENK, 1929. *Z. techn. Phys.* **10**, 57.
SCRASE, 1931. London, Meteor. Office, *Geophys. Mem.* no. 52.
SHEPPARD, 1939. *J. Sci. Instrum.* **9**, 218–21.
SHERLOCK and STOUT, 1931. *Engr. Res. Bull. Univ. Michigan*, no. 20.
SHERLOCK and STOUT, 1937. *J. Aero. Sci.* **5**, 53–61.
SIMMONS, 1934. *Proc. Roy. Soc.* A, **145**.
SIMMONS, 1938. Reported in GOLDSTEIN, 1938, Vol. 1, Sect. vi.
SIMMONS and BAILEY, 1926. *Rep. Memor. Aero. Res. Comm., Lond.*, no. 1019.
SPILHAUS, 1934. *Mass. Inst. Techn. Prof. Notes*, no. 7.
TAYLOR, 1927. *Quart. J. R. Met. Soc.* **53**, 201.
TOWNEND, 1934. *Proc. Roy. Soc.* A, **145**, 180–211.
WHEWELL, 1835. *Rep. Brit. Ass.* **2**, 32.
WOLLASTON, 1829. *Philos. Trans.* **119**, 133.

LLOYD'S SINGLE-MIRROR INTERFERENCE FRINGES

By P. F. TITCHMARSH, B.Sc.,
Sheffield

Communicated by J. H. Awbery ; *received* 8 *February* 1941 ;
in revised form 31 *March* 1941

ABSTRACT. Half of the central fringe of Lloyd's single-mirror system can be seen directly if the fringes are observed in a plane passing through the edge of the mirrors, and if the mirror is of ordinary, not black, glass and has a cleanly cut edge which is illuminated by light shining on to the opposite edge.

The use of mica to displace the fringe system is not valid, since the central fringe will be displaced more for blue light than for red. This dispersion of the central fringes for the different colours may produce a displaced white-light system approximately symmetrical about a black fringe, but this black fringe will not be produced by the coincidence of the displaced central fringes. It is equally possible to find a thickness of mica which gives a displaced system symmetrical about a white fringe.

The phase change on reflection at a less dense medium at grazing incidence is shown to be in the region of π, since a system with a black central fringe is formed when the mirror is immersed in carbon bisulphide.

§ 1. INTRODUCTION

Lloyd's method of producing interference fringes by the use of a single mirror is of interest for two reasons: firstly, because it is the simplest method of obtaining interference fringes, requiring no accurately made apparatus and little skill in adjustment; secondly, because it affords a striking demonstration that there is a phase change of π when light travelling in air is reflected at the surface of a denser medium.

The object of this paper is to clear up some discrepancies which I have found between Lloyd's account, the accounts of the experiment given in text-books, and my own observations. I am mainly concerned with the two following points:—

(*a*) The fact of the phase change on reflection can be established by *direct* observation of half of the central fringe, which can be seen to be black.

(*b*) The displacement of the fringe system by means of a mica plate is valueless, since observation of the displaced system can yield no evidence of the nature of the central fringe of the undisplaced system.

I have also investigated the nature of the interference system produced by reflection from an optically less dense medium.

Lloyd's account

The following is part of Lloyd's own account, taken from a paper which was read to the Royal Irish Academy on January 27th, 1834, and published in the *Transactions* of that body, vol. 17.

After a theoretical discussion Lloyd says:—

" In order to submit these results to the test of trial, I employed the apparatus consisting of two movable plates, which is of so much use in experiments of interference. The plates being closed, so as to form a narrow horizontal aperture, the flame of a lamp was placed behind ; and the light thus diverging from the aperture was received, at the distance of about three feet, on a piece of black glass truly polished, and also horizontal. The reflector was then adjusted, so that its plane might pass a little below the aperture ; or, in other words, that the light might be incident on it at an angle of nearly 90°. It is evident that the light thus obliquely reflected will meet the direct light diverging from the aperture under a very small angle, and with a difference in the lengths of their paths which is capable of indefinite diminution. The two lights are therefore in a condition to interfere : and I found, accordingly, that when they were received upon an eyepiece, placed at a short distance from the reflector, a very beautiful system of bands was visible, in every respect similar to one-half of the system formed by the two mirrors in Fresnel's experiment.

" The first band was a *bright* one and *colourless*. This was succeeded by a very sharply defined black band ; then followed by a coloured band and so on alternately. Under favourable circumstances I could easily count seven alternations : the breadth of the bands being, as far as the eye could judge, the same throughout the series, and increasing with the obliquity of the reflected beam. The first dark band was of *intense blackness* ; but the darkness of the succeeding bands was less intense, as they were of higher orders ; and after three or four orders they were completely obliterated by the closing in of the bright bands. At the same time the coloration of the bright bands increased with the order of the band ; until, after six or seven alternations, the colours of the different orders became superimposed, and the bands were thus lost in a diffused light of nearly uniform intensity. All these circumstances are similar to those observed in Fresnel's experiment and correspond exactly with the results of theory.

" These bands are most perfectly defined when the eyepiece is close to the reflector. Their breadth and coloration increased with the distance of the eyepiece, but remained of a finite and very sensible magnitude when the latter was brought into actual contact with the edge, a circumstance which distinguishes them altogether from the diffracted fringes formed on the boundary of the shadow."

He then discusses the problem of the phase change on reflection, and continues:—

" The present case of interference seems to support this view " (i.e. that the phase change occurs on reflection at the denser medium). " It follows from theory that if the light undergoes no change of phase by reflection, the distances of the successive dark fringes from the edge of the shadow will be as the odd numbers 1, 3, 5, etc. ; so that the distance of the first dark band from the edge will be half the interval between each succeeding pair of dark bands. But it appears, on the contrary, that the distance is, as far as the eye can judge, exactly equal to the succeeding intervals ; or that the bands are all shifted from the edge by the amount of *half an interval* "

Figure 1 (to which reference will be made later) is copied, with one or two minor changes, from this paper.

Text-book accounts

The following account, typical of that in various text-books, is taken from *Light for Students*, by E. Edser:—

"Dr. Lloyd obtained fringes by causing light from a narrow slit to be split up into two pencils, one direct, and the other reflected from a polished black glass surface. In this case the wave sources are the illuminated slit itself (A, figure 2), and the virtual image, B, of the slit obtained by reflection at the glass surface. As will be seen from the diagram, the point M will not be illuminated by reflected light, and the central fringe consequently is not formed; it can, however, be brought into view by placing a thin film of mica or glass in the path of the directly transmitted pencil. Owing to the retardation of the waves which traverse the film, the position of the central fringe is displaced towards P. If the film is of suitable thickness, the position of the central fringe will be situated on that part of the screen illuminated by the reflected waves.

"Dr. Lloyd observed that the central fringe is black, instead of being white as in most other cases. On either side of the central black band is a white fringe, the rest of the fringes being coloured. This indicates that, *when light is reflected from an optically denser medium, the phase of the reflected waves differs by π from that of the incident waves.*

The direct observation of the central fringe

The most important difference between the two accounts is seen by comparing the two diagrams. Lloyd considered the fringe system formed in a plane passing through the edge of the mirror, whereas the text-books consider a plane of observation at some distance from the mirror. In the second case it is impossible to see the central portions of the fringe system directly, but in the first case it should theoretically be possible to see the whole of the central portion up to the point where the path difference between the direct and reflected rays approaches zero.

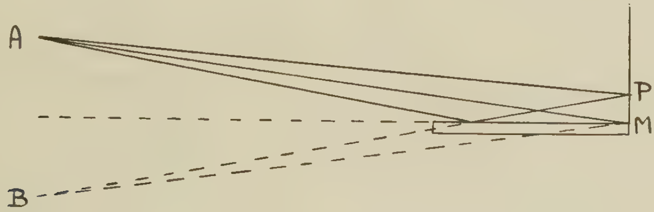


Figure 1.

Actual observation at once shows Lloyd's to be the more practical arrangement, it being, in the words of Wood's *Physical Optics* "... almost impossible to miss finding the fringes at the first attempt". With the apparatus arranged as in the text-book diagram, the central fringes are cut out, and these, especially in the case of white light, are the most sharply defined fringes. With Lloyd's arrangement the fringes will be seen even if no great care has been taken to make the slit parallel to the plane of the mirror. This adjustment can be made by observing the slit and its image and moving either the slit or the mirror until the two sources appear parallel. When this observation is impossible, e.g. when

a camera is fixed in position behind the mirror, the adjustment is easily made if a piece of white card is placed with one edge in contact with the face of the mirror; an image of the slit is then thrown on to this card by a convex lens and the slit is adjusted until the image appears parallel to the edge of the card.

Now if we observe the fringe system in a plane passing through the edge of the mirror, what shall we expect to see? Lloyd says " . . . a very beautiful system of bands was visible, in every respect similar to one-half of the system formed by the two mirrors in Fresnel's experiment", except that " . . . the bands are all shifted from the edge by the amount of half an interval". This statement is incorrect. The fringe system produced by Fresnel's mirrors with a slit illuminated by white light is symmetrical about a white central fringe.

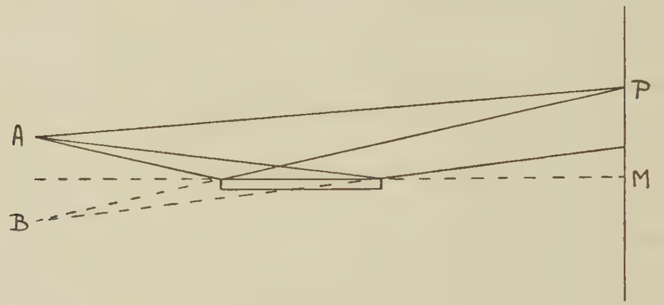


Figure 2.

Only the central point is truly achromatic, the dark fringes on either side showing some colour. This system may be thought of as being formed by the superposition of a large number of monochromatic fringe systems, each having a bright fringe at the same central point (the point of no path difference between the interfering rays). Lloyd's system is formed in a similar way, except that a phase change of π is introduced into one of the rays. This means that at the point of no path-difference each monochromatic fringe system will have a dark fringe, and the white-light system will be symmetrical about a dark central fringe, with a slightly coloured bright fringe on either side (only half of this system being visible). The same result is obtained by supposing the central fringe of each *monochromatic* system (and with it every other fringe in the system) to be shifted by half a fringe-width from the position in Fresnel's system, but this is *not* the same as supposing the whole of Fresnel's white-light system to be shifted through half a fringe-width. Lloyd's and Fresnel's white-light systems are, in fact, exactly complementary. The mistake seems to have survived in a number of text-books; e. g. Wood states that " Dr. Lloyd found that the centre of the system did not lie on the plane of the surface, as might be expected, but was displaced by the width of half a fringe". Such a statement is misleading, and can only be said to be true if it is explicitly stated to apply to a monochromatic fringe system.

Thus, if we observe in a plane through the edge of the mirror we may expect

to see one-half of a fringe system which has at its central point, which is the only truly achromatic point when white light is used, a dark fringe. It is the existence of this central dark fringe which shows that there is a phase change of π on reflection at the glass surface.

Since we can only see one-half of the system we can only hope to see one-half of this central fringe. The text-books seem to dismiss even this possibility, and to rely on moving the whole system by means of a mica plate. But with Lloyd's arrangement it should be possible, if the apparatus conforms sufficiently well to the theoretical conditions. The most critical condition is that the plane of observation shall be focused exactly on the mirror edge, for any imperfection in this focusing will produce blurring in exactly that part of the field of view that we wish to examine. For the focusing to be exact it is necessary that the mirror edge shall be very accurately cut. I have found that such an edge can be obtained from ordinary picture glass ($^1/_{12}$-inch thick). When such glass is cut by scratching along one side with a diamond and breaking off, one edge is ruined by the diamond, but the edge on the opposite side of the break appears to be almost perfect, i. e. very few imperfections are visible under the magnifications used to view the fringes. Although such glass has not an optically worked surface, it is sufficiently plane to produce clearly defined fringes; irregularities may be apparent in the fringe system, but a reasonably good result can be obtained by trying a number of pieces of glass. Any piece of glass which has had its surface worked after cutting will not have a perfect edge. There seems to be no advantage in using black glass, as any secondary reflections are of negligible intensity.

When the fringe system is observed, the half of the dark central fringe may not at once be apparent, for it may not be possible to distinguish this half-fringe from the edge of the mirror itself upon which it lies. Since the half-fringe is dark, the edge of the mirror must be illuminated in some way to provide a contrast. This is best done by throwing a beam of light on to the edge of the mirror which is pointing towards the slit. This light is totally reflected a number of times within the glass and emerges at the edge which is being observed. The arrangement is shown in figure 3. The light must be projected from behind the mirror, otherwise some of it will be reflected from the front surface and will illuminate the part of the field of view to be occupied by the fringe system. For a similar reason, a screen, X, must prevent light from entering the back of the mirror. If the incident beam makes a small angle with the plane of the mirror, then the edge under observation will appear uniformly illuminated. The intensity of the illumination can be adjusted by a slight rotation of the mirror M_2. When the mirror edge is suitably illuminated, the half of the dark central fringe can be clearly seen. Evidently Lloyd missed seeing this because he used black glass for the mirror, and because the magnifications he used were, presumably, comparatively small. Thus he emphasizes that " the first dark fringe was of *intense blackness* ", but under a high magnification the coloration of this band is quite clear.

The half of the central dark fringe is shown in photographs 1, 2 and 3.

Photograph 1 was taken with white light illuminating the slit. The contrast of light and dark in this fringe system is rather greater in the photograph than when it was observed by eye. This is because the plates used (Agfa Isopan, Agfa Isochrome and Ilford H.P.2) were more sensitive to the red end of the spectrum.

Photograph 2 was taken with yellow light (from a spectroscope) and with the slit closer to the plane of the mirror than in photograph 1.

Photograph 3 was taken with green light and with the slit further from the plane of the mirror than in photographs 1 and 2.

The half-fringe appears comparatively wider in photograph 1 than in photographs 2 and 3 ; this is because the central dark fringe in the white-light system is broader than the " dark " fringes on either side, which are actually coloured at the edges. In photograph 3 the dark fringes are wider compared with the bright fringes owing to a shorter exposure time.

When the half-fringe is seen for the first time it is natural to cast doubts on its genuineness. Is it a true interference fringe? Might it not be a shadow, especially when the slit is very close to the plane of the mirror and the fringes are very wide? If the edge of the mirror is cleanly cut, then a shadow is seen only when the slit is behind the plane of the mirror and interference fringes are replaced by the diffraction fringes from the far edge. If the mirror edge is damaged then the dark line appears wider than half a fringe at some places, the added width being shadow and not interference fringe. In all other cases a half-fringe appears in the position required by theory. It is the agreement of observation with theory over a wide range of fringe-widths and with a number of different pieces of glass that gives the conclusive proof of the nature of the dark line.

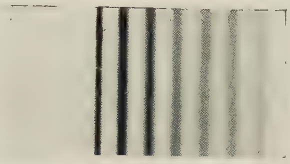
§2. THE DISPLACEMENT OF THE FRINGE SYSTEM BY A MICA PLATE

According to the text-books it is possible to displace the central portion of the fringe system outwards from the mirror by causing the direct rays to pass through a plate of mica. To test this I used the arrangement shown in figure 3. The mica plate was about 2 cm. $\times \frac{1}{2}$ cm. and of a suitable thickness (about a thousandth of a centimetre) obtained by splitting. The gap between the edge of the plate and the face of the mirror was about 1 mm. and could be accurately controlled by a screw.

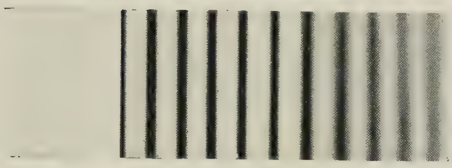
The field of view, with white light illuminating the slit, was as shown in the upper section of photograph 4.

It will be noticed that both displaced and undisplaced fringe systems can be seen; this is possible because the gap between the mica plate and the mirror allows light, both direct and reflected, to fall on a small region of the field of view near to the mirror edge. The half-fringe appears wider than it should in photographs 4 and 5. This is because an optically worked plate was used, the edge of which had been slightly worn away by polishing.

Photograph
1
White

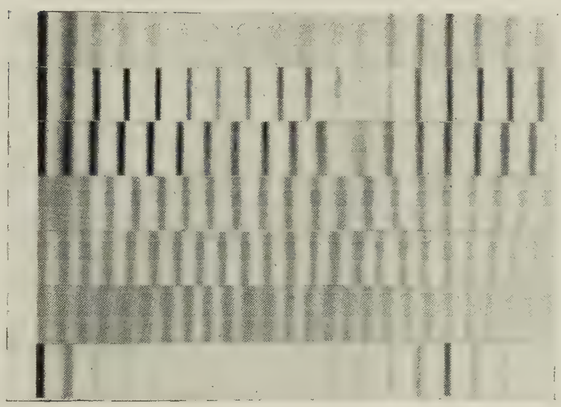


2
Yellow



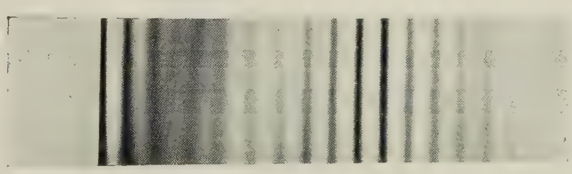
3
Green

4
White
Red
Yellow
Green
Blue
Violet
White



$t =$
0·0013 cm.
6200 A.
5700 A.
5200 A.
4500 A.
4200 A.

5
White



$t =$
0·0010 cm.

Magnifications:—Photographs 1, 2, 3, 5, ×70.
Photograph 4, ×85.

To face page 396

A casual glance might seem to show that the displaced system is as predicted, i.e. with a central black fringe; but careful study revealed two facts:—

(*a*) The central " black " fringe was not perfectly colourless, having a very faint red tinge on the side nearest to the mirror, and a definite green coloration on the further side.

(*b*) The fringe system was not perfectly symmetrical about this central fringe, the fringes on the side away from the mirror being distinctly more coloured than the corresponding fringes on the side nearer to the mirror. (All the fringes except the central one were coloured red on the side towards the central fringe and green on the opposite side.)

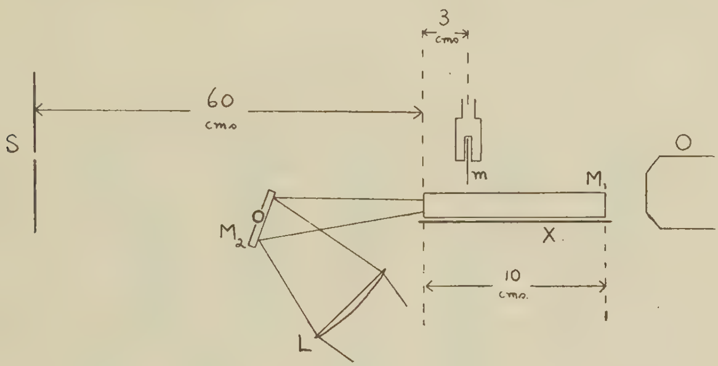


Figure 3

L. Lens focusing light on to the edge of M_1.
M_1. Lloyd's mirror.
M_2. Mirror controlling the intensity of the illumination of M_1.
m. Mica plate.
O. Camera lens. A ⅔-inch microscopic objective.
S. Slit.
X. Screen to prevent light from entering the back of M_1.

It did not require any critical adjustment of the apparatus to bring the displaced system into view. The displacement of the central fringe from the mirror edge appeared to be a constant number of fringe-widths. Thus, as the fringe-width was increased, by changing the inclination of the mirror, the displacement of the system was also increased, and *vice versa*. The limits of fringe-width within which the displaced system could be seen could be increased up to a certain point by increasing the gap between the mica plate and the mirror. Thus it was possible to observe the displaced system in a number of different positions. I observed that the divergencies from prediction were present in all such positions. From this I concluded that these divergencies were of a fundamental nature, and were not due to casual inaccuracies in the apparatus (for in the different positions the reflected rays would meet different parts of the mirror surface, and the direct rays would pass through different parts of the mica plate).

The solution to the problem of this divergence from prediction lies in the

consideration of the dispersive power of the mica, which many of the text-books seem to have ignored.

Let the thickness of the mica plate be t and let its refractive index, for a certain wave-length λ, be μ.

Then the path difference produced by the introduction of the mica plate is $t(\mu-1)$.

If interference is obtained from sources distant d apart and the fringes are observed in a plane distant l from the sources, then the path difference at a point in this plane at a distance x from the centre of the system is xd/l.

Thus when the mica is introduced, the point of zero path-difference is displaced x outwards from the mirror edge, where

$$xd/l = t(\mu-1).$$

Thus $$x = lt(\mu-1)/d.$$

Evidently this displacement depends upon λ, since μ varies with λ. Thus the central black fringes of the fringe systems due to the various wave-lengths will not all be displaced to the same point, i.e. there will not be a displaced system which is perfectly symmetrical about a black fringe.

It may be, however, that the effect of the dispersion on the displaced system is negligible. Thus it is necessary to examine the displacement of the fringes in rather more detail.

As no figures for the dispersive power of mica are available, let us assume that the variation of μ with λ for mica is of the same nature as that for crown glass.* We can draw (from details in Kaye and Laby's *Tables*) a graph of $\mu-1$ against λ for crown glass.

From this graph we can construct the following table, showing the relative displacements of the central dark fringes of the systems due to the various wave-lengths:—

Fringe-width (in.)	λ (A.)	Colour	$\mu-1$	Displacement factor	Actual displacement (in.)	Relative displacement (in.)
1·0	6600	Red	0·5143	1·000	12·00	0·00
0·9	5940	Yellow	0·5167	1·004	12·05	0·05
0·8	5280	Green	0·5200	1·011	12·13	0·13
0·7	4620	Blue	0·5253	1·021	12·25	0·25
0·6	3960	Violet	0·5332	1·036	12·43	0·43

Five wave-lengths are chosen, so that the fringe-widths in the resulting interference systems may be represented in a diagram by the numbers in the first column. The corresponding values of $\mu-1$ are read off from the graph. We now calculate for each wave-length a "displacement factor" showing how much the central fringe is displaced compared with the displacement for $\lambda = 6600$ A. Since displacements are proportional to $\mu-1$, the factor for $\lambda = 5940$ A. is

* It has since been pointed out to me that some figures for the refractive index of mica in the red half of the spectrum are given in Landolt-Bornstein's tables. These figures yield values of the "displacement factor" almost identical with those calculated for crown glass, though the actual refractive indices are different.

$\frac{0 \cdot 5167}{0 \cdot 5143} = 1 \cdot 004$. Now suppose that the mica displaces the central fringe of the 6600 A. system through 12 fringe-widths (corresponding approximately to the piece of mica used in the experiment), the actual displacement in our diagram will be 12 in. The actual displacement for the central fringe of the 5940 A. system will be $1 \cdot 004 \times 12$ in. $= 12 \cdot 05$ in., i.e. 0·05 in. further from the mirror edge than the central dark fringe of the red system. It is from the first and last columns of this table that figure 4 is constructed.

Each fringe system is represented by light and dark spaces on a strip of graph paper. The line CC shows the position of the central dark fringe of the 6600 A.

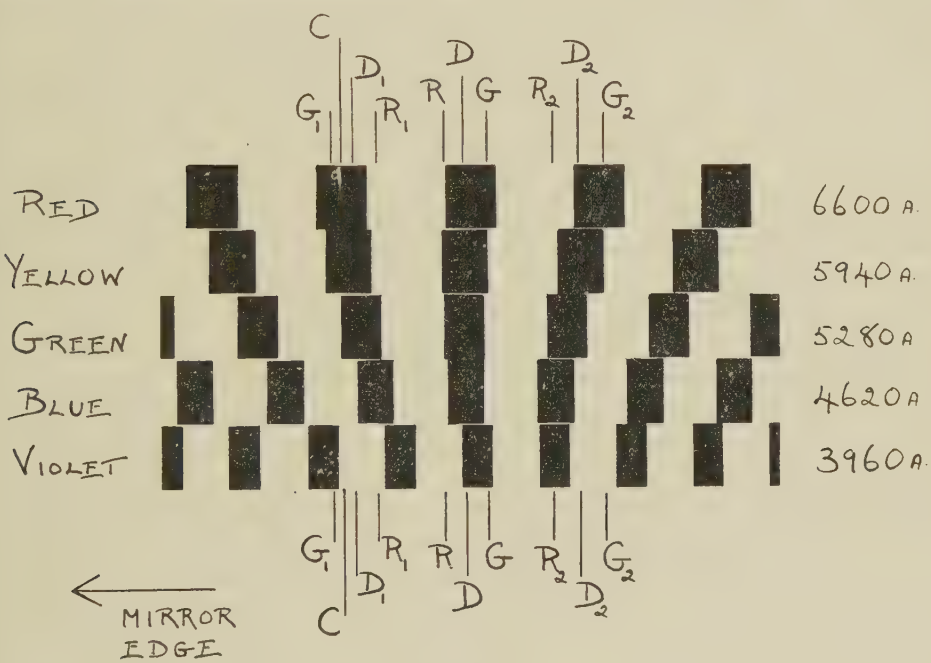


Figure 4.

system which has been displaced 12 fringe-widths from the edge of the mirror (supposed to be to the left of the diagram). The central fringe of the 5940 A. system is displaced 0·05 in. to the right of this, and the other systems are similarly displaced.

We can determine the resulting illumination at any point by considering a line across the diagram parallel to CC. (In doing this it must be remembered that there is actually no sudden change from zero to maximum illumination in any fringe system.) It is clear that there is only one point at which the illumination will be closely approaching zero, and that is at D. To the left of this there will be a very slight preponderance of red at R; to the right there will be a definite preponderance of green at G. The next dark fringe, in order of distinctness,

will have its centre at D_1, where the illumination is mainly due to the violet, to which the eye is not very sensitive. On either side of D_1 there will be definite coloration, red at R_1 and green at G_1. The next fringe, with its centre at D_2, will have still more coloration, R_2 and G_2 being a more definite red and green than R_1 and G_1.

These conclusions agree with the observations previously described; but it would be more satisfactory to have some direct evidence that the mica is actually dispersing the centres of the fringe systems to the extent that we have assumed. With this object in view I took a set of photographs of the displaced system, including in each case the mirror edge, using in turn white, red, yellow, green, blue and violet light to illuminate the slit. Since the results had to be compared, careful measurements were taken of the white-light fringes both before and after the exposures, and the results were rejected until the measurements showed that there had been no change. By cutting narrow strips from these photographs I obtained seven strips (five colours and an initial and a final white exposure) similar to the strips of graph paper from which figure 4 was constructed. These were stuck on to a common base with the photographs of the mirror edge in line. The result should resemble figure 4, and should give us a measure of the dispersive power of the mica. If that dispersive power is very low it may be that we shall find a set of dark fringes which are in a good straight line except for slight deviations in the blue and violet.

The composite picture is shown in photograph 4. Its resemblance to figure 4 is quite clear. The set of dark fringes which produces the darkest fringe of the white-light system cannot be the set of central fringes which has been dispersed by the mica, for of these the green is nearest to the mirror edge. So the central fringes must form the previous set, which produce the second darkest fringe of the white-light system. Thus the mica is shown to disperse the centres of the fringe system to the extent required by the theoretical discussion.

The most important fact revealed by the theoretical discussion is that the nearly black central fringe is not produced by the coincidence of the central dark fringes of the different systems, but by the fringes next to the right, i.e. away from the mirror, from the central fringes. Thus when we observe the displaced system and see a nearly black fringe at its centre, we are not observing the central fringe of the original system in its displaced position. This raises the question as to whether the observation of the displaced system can give any evidence as to the nature of the central fringe of the undisplaced system.

The piece of mica used in the experiment has dispersed the centres of the fringe systems so that the next set of dark fringes are nearly coincident. If a thinner piece of mica were used, the dispersion of the centres of the fringe systems would be less ; and, for a certain thickness, this would cause the *bright* fringes next to the central dark fringes to coincide, i.e. the white-light fringe system would be symmetrical about a central *white* fringe. Such a system is shown in photograph 5, in which there are two dark lines clearly distinguishable from the other fringes. Actually the right-hand fringe of this pair is appreciably darker

than the other, but when viewed by eye the system was clearly symmetrical about a white central fringe rather than about a black one. A slightly thinner piece of mica would give more perfect symmetry.

It is clear, both from the theoretical discussion and from the photographs, that by varying the thickness of the mica any type of symmetry can be produced in the displaced system. Thus I conclude that the presence of a nearly black fringe, distinguishable from all other dark fringes, approximately at the centre of the displaced system, is no evidence that the central fringe of the undisplaced system is black; and that therefore the only evidence as to the nature of the central fringe lies in its direct observation.

§ 3. THE FRINGE SYSTEM PRODUCED BY REFLECTION AT A LESS DENSE MEDIUM

The existence of the dark central fringe in Lloyd's single-mirror interference system shows that there is a phase change of π when light is reflected from an optically denser medium. What, then, would be the nature of a similar interference system produced by reflection from an optically less dense medium?

To investigate this question I immersed the mirror in carbon bisulphide. The liquid was contained in a tank of which the end plates, through which the light passed, were optically worked. By the "silvery" appearance of the liquid-glass surface when seen by light incident at large angles, it was clear that total reflection was taking place, i.e. that the liquid was the optically denser medium. When the fringe system was observed it appeared to be identical with the system produced in the ordinary case. The fringe-pattern changed during observation owing to convection currents in the liquid. The rate of change was deliberately increased by agitating the liquid. The width of the fringes changed considerably when the tank was moved, so as to alter the inclination of the end plates to the rays of light. But in all cases the half of the dark central fringe could be seen on the edge of the mirror. Thus when light is reflected at grazing incidence from the surface of a less dense medium there is also a phase change of π.

This result agrees with the theoretical treatment given in Preston's *Theory of Light* (5th edition), Article 214, where it is deduced that:

If i = angle of incidence of the light in the denser medium,

μ = index of refraction from the less dense to the denser medium,

δ = change of phase on reflection of light polarized in the plane of incidence (i.e. with electric vibrations perpendicular to the plane of incidence), and

δ' = change of phase on reflection of light polarized perpendicular to the plane of incidence (i.e. with electric vibrations in the plane of incidence),

then

$$\cos\delta = \frac{\mu^2+1-2\mu^2\sin^2 i}{\mu^2-1}; \quad \cos\delta' = \frac{\mu^2+1-(\mu^4+1)\sin^2 i}{(\mu^4-1)\sin^2 i-(\mu^2-1)}.$$

(Actually the value given in the book for $\cos \delta'$ is minus the value given here. This is due to the fact that in the book the positive directions of the incident and reflected vibrations in the plane of incidence are opposed when i is large—see footnote on p. 393.)

When $i=90°$, $\cos \delta = \cos \delta' = -1$ and $\delta = \delta' = \pi$; so that at grazing incidence both components suffer a phase change of π on reflection.

§ 4. A COMMENT ON THE EXACT SIGNIFICANCE OF THE RESULTS

Lloyd's experiment is normally regarded as a demonstration that the phase change on reflection at a denser medium is π *rather than zero*. From this aspect the result is quite definite—there is a phase change of π. But if we are to ask, as we must in the experiment on reflection at a less dense medium, " Is the phase change exactly equal to π? ", then the answer is much less definite. If the distance from the mirror edge to the centre of the first bright fringe is f, and the fringe-width is F, then the phase change on reflection is $2\pi \cdot f/F$. Thus the accuracy of the estimated value of the phase change depends on the measurement of the two small distances f and F, and so it cannot be very high. In the experiments described here the accuracy is still further reduced by the fact that the plate which has the accurately cut edge, from which measurements can be made, does not give a perfectly uniform fringe-width. Measurement of the negative of photograph 2 shows that the phase change on reflection at a denser medium must lie between 170° and 190°.

In the case of reflection at a less dense medium no measurements could be made owing to the fluctuation of the fringe-pattern. The estimation that the phase change was equal to π depended on the observation that there was half of a dark fringe on the mirror edge ; this dark line would differ appreciably from a half-fringe only if the phase change lay outside limits of about 150° and 210°. If we take the refractive indices of the mirror and of carbon bisulphide to be 1·52 and 1·63, and if $i=88°$, then $\delta = 168° \, 54'$, $\delta' = 170° \, 22'$, and the fringe-width is about one-third of that in photograph 3. Thus δ and δ' will not produce any visible alteration of the fringe system in this case. We might observe an alteration if we used a liquid which had a refractive index only slightly greater than that of the glass. Thus if the refractive indices of the mirror and liquid were 1·520 and 1·522, the critical angle would be 87°, and for $i=89°$, $\delta = 140° \, 0'$, $\delta' = 140° \, 8'$, and the fringe-width about two-thirds of that in photograph 3. In such a case, if care were taken to eliminate convection currents in the liquid we might hope to obtain photographic evidence of a phase change differing from π.

§ 5. ACKNOWLEDGEMENTS

In conclusion I should like to thank Dr. A. W. Barton and Dr. G. K. T. Conn for their advice on the preparation of this paper, and M. W. Johnson and D. O. Malthouse, of the Sixth Form of King Edward VII School, Sheffield, for their assistance with the photography.

DETERMINATION OF THE SURFACE TENSION AND ANGLE OF CONTACT OF A LIQUID

By A. E. BATE, M.Sc., Ph.D., F.Inst.P.,
Northern Polytechnic, London, N.7

Received 13 *January* 1941 ; *in revised form* 1 *April* 1941

ABSTRACT. The theory is based on the large sessile drop. The expressions

$$h^2+d^2=4a^2(1+2a/3r)$$

and $$\cos\theta=(d^2-h^2)/(d^2+h^2)$$

are derived, in which h is the height of a drop of the liquid and d is the depth of a bubble formed in the liquid beneath a horizontal plane ; a^2 is $T/g\rho$, T being the surface tension required, and r the horizontal radii of the drop and bubble, assumed to be the same and at least 6 cm. in length. The first expression is independent of the angle of contact θ, and the second is independent of the surface tension and density ρ of the liquid. Actually the latter expression is true for bubbles of infinite extent only, but the error incurred in using it for finite bubbles of the radius mentioned is a fraction of 1 per cent of $\cos\theta$.

The methods of developing the drop and bubble and of measuring h and d are described ; preliminary results indicate that benzene and water have definite angles of contact with plate glass.

§ 1. INTRODUCTION

Several of the methods of determining the surface tension of a liquid depend on a knowledge of the angle of contact between the surface of the liquid and the containing vessel, but the actual methods of measuring the angles of contact are not always reliable.

In this paper a method based on the theory of the sessile drop is presented by which the surface tension of a liquid and its angle of contact can be measured with some degree of certainty.

§ 2. THEORY

The method is based in principle on the sessile drop, and the procedure consists first in developing a *bubble* of large extent in the liquid beneath a plane of plate glass ; next, a *drop* of similar extent is formed on the plate glass. In each case the plane of glass is horizontal, and the depth of the bubble below the plate, and the height of the drop above it, are measured. The appearance for a liquid whose angle of contact is about 30° will be as in figure 1, which gives the maximum cross-sections of the drop and bubble respectively.

The faces both of bubble and drop will be horizontal. At the point P, where the plane of maximum horizontal cross-section cuts the bubble, let R and r

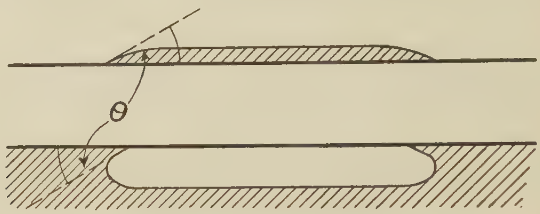


Figure 1.

represent the radii of curvature of the horizontal and vertical sections respectively. The pressure p inside the bubble will be given by

$$p=2T\left(\frac{1}{r}+\frac{1}{R}\right), \qquad \text{......(1)}$$

and when R is exceedingly large,

$$p=2T\left(\frac{1}{r}\right). \qquad \text{......(2)}$$

These equations indicate that the pressure diminishes as the bubble increases in volume, i.e. the bubble is slightly shallower when *very* large than when large, the diameter of the former being of the order 100 cm., whereas the latter may be about 4 cm. in diameter (Burdon, 1940).

From elementary principles it may be shown, for a very large drop of height h, that

$$T-T\cos\theta=h^2g\rho/2, \qquad \text{......(3)}$$

and for a bubble of similar extent, but of depth d, that

$$T+T\cos\theta=d^2g\rho/2. \qquad \text{......(4)}$$

By adding equations (3) and (4), the expression

$$2T=(h^2+d^2)g\rho/2 \qquad \text{......(5)}$$

is obtained, and this is independent of θ ; by subtraction,

$$\cos\theta=(d^2-h^2)g\rho/2T=(d^2-h^2)/(d^2+h^2), \qquad \text{......(6)}$$

and this is independent of T and ρ. Thus, by measuring h and d, T and θ can be found.

If the contact angle is zero, expression (3) shows that it is impossible to form a drop, whereas (4) indicates that it is possible to form a bubble, and that it will be complete. Let D stand for the depth of such a bubble; then, if it is of infinite extent, expression (4) becomes $2T=D^2g\rho/2$, i.e. from (5),

$$D^2=h^2+d^2. \qquad \text{......(7)}$$

In practice, bubbles of small extent are used, and a correction is required for the slightly greater depths of bubble and drop which occur, as indicated above.

Ferguson (1913) has given such a correction ; he has shown that, when a bubble is large and, therefore, has a plane face,

$$k^2 = 4a^2 \sin^2 \theta/2 + 8a^3(1 - \cos^3 \theta/2)/3r, \qquad \ldots\ldots(8)$$

in which $a^2 = T/g\rho$,

k = total depth of bubble,

and θ = angle of contact.

Applying this to a finite bubble of depth d_1 and to the corresponding drop of height h_1, we have

$$h_1^2 = 4a^2 \sin^2 \theta/2 + 8a^3(1 - \cos^3 \theta/2)/3r \qquad \ldots\ldots(9)$$

and

$$d_1^2 = 4a^2 \sin^2 (180 - \theta)/2 + 8a^3(1 - \cos^3 \overline{180 - \theta/2)}/3r$$
$$= 4a^2 \cos^2 \theta/2 + 8a^3(1 - \sin^3 \theta/2)/3r. \qquad \ldots\ldots(10)$$

Adding (9) and (10) gives

$$h_1^2 + d_1^2 = 4a^2 \left[1 + \frac{2a}{3r}(2 - \overline{\sin^3 \theta/2 + \cos^3 \theta/2)} \right]. \qquad \ldots\ldots(11)$$

In most cases θ is less than 10°, and as $\sin^3 5° + \cos^3 5° = 0{\cdot}99$, the value of this quantity in the small term may be taken as unity with no appreciable error, and (11) becomes

$$h_1^2 + d_1^2 = 4a^2(1 + 2a/3r). \qquad \ldots\ldots(12)$$

If, now, $\theta = 0$, from (9) and (10)

$$h_0^2 = 4a^2(1 + 2a/3r) \quad \text{and} \quad d_0^2 = 0.$$

Again, for $\theta = 180°$,

$$h_{180}^2 = 0 \quad \text{and} \quad d_{180}^2 = 4a^2(1 + 2a/3r),$$

and thus the height of a drop on a plate of contact angle 180° is the same as the depth of a bubble in the same liquid beneath a plate of contact angle 0°. Denoting this particular value by H_1, then

$$H_1^2 = 4a^2(1 + 2a/3r). \qquad \ldots\ldots(13)$$

At first sight this appears to imply that $H_1^2 = h_1^2 + d_1^2$, but it must be emphasized that this is only approximately true, and that when the angle is small. An example will make this clear. Consider a liquid in contact with a plate, the angle of contact

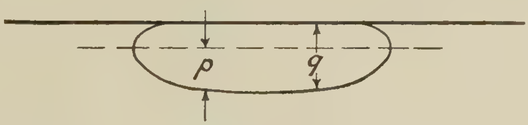


Figure 2.

being 90°. In this case the depth of the bubble will be the same as the height of the drop : if, now, a bubble be formed in the same liquid beneath a plane which has a zero contact angle, the horizontal plane of maximum area will divide the bubble as shown in figure 2, the lower portion being deeper than the upper. The ratio for an infinite bubble is $p/q = 1/\sqrt{2}$. Thus a bubble and the corresponding drop do not together constitute the complete bubble, and H_1 must be regarded as the height of a fictitious bubble.

If the bubble becomes very large, r is very large, and the small term in the expressions (8) to (12) becomes negligible, and $H^2 = 4a^2$ in (13). When, however, the radius is 6 cm., the small term may be of the order 5 per cent, so that, following Ferguson, $2a$ is replaced by H as an approximate value in (13), which becomes

$$H^2 = 4a^2(1 + H/3r), \qquad \ldots\ldots(14)$$

from which $4a^2$ is obtained. The new value of $2a$ is inserted in (13) and $4a^2$ again calculated; this is repeated until the value obtained equals that substituted. This is the method of successive approximations.

The value finally obtained is used in the expression $a^2 = T/g\rho$ to determine T. The same procedure is adopted for liquids with appreciable angles of contact, except that expression (11) must be used, with an approximate value of θ substituted. This merely gives (14) a different coefficient in the last term.

This approximate value θ of the angle of contact is obtained from expression (6), in which the values d_1, h_1 are substituted for d and h respectively; denoting this value of the angle by θ_1, the expression becomes

$$\cos\theta_1 = \frac{d_1^2 - h_1^2}{d_1^2 + h_1^2} = \frac{4a^2[(\cos^2\theta/2 - \sin^2\theta/2) + 2a(\cos^3\theta/2 - \sin^3\theta/2)/3r]}{4a^2[(\cos^2\theta/2 + \sin^2\theta/2) + 2a(2 - \cos^3\theta/2 - \sin^3\theta/2)/3r]}$$

$$= \frac{\cos\theta + 2a(\cos^3\theta/2 - \sin^3\theta/2)/3r}{1 + 2a(2 - \cos^3\theta/2 - \sin^3\theta/2)/3r}, \qquad \ldots\ldots(15)$$

after substituting for d_1 and h_1 from (9) and (10).

Actually, however, the difference between θ and θ_1 does not exceed a fraction of a degree, for $2a$ is less than $0{\cdot}5$ cm. in all cases, and r is made as near 6 cm. as possible, so that $2a/3r$ is about $1/40$ or less. Write expression (15) as

$$\cos\theta_1 = \frac{\cos\theta + A\,.\,2a/3r}{1 + B\,.\,2a/3r}, \qquad \ldots\ldots(16)$$

in which A stands for

$$\cos^3\theta/2 - \sin^3\theta/2$$

and B stands for

$$2 - \cos^3\theta/2 - \sin^3\theta/2,$$

and this expression becomes

$$\cos\theta_1 = \frac{\cos\theta + A/40}{1 + B/40}.$$

This would make the actual value θ equal the approximate value θ_1 if $A/B = \cos\theta$.

The following table gives the values of $\cos\theta_1$ for several values of θ :—

θ (degrees)	$\cos\theta$	$\cos\theta/2$	$\cos^3\theta/2$	$\sin\theta/2$	$\sin^3\theta/2$	A	B	$\cos\theta_1$
0	1	1	1	0	0	0	0	1
30	0·866	0·966	0·90	0·259	0·02	0·88	1·08	0·865
60	0·5	0·866	0·65	0·5	0·13	0·52	1·22	0·498
90	0	0·707	0·35	0·707	0·35	0	1·3	0
120	−0·5	0·5	0·13	0·866	0·65	−0·52	1·22	−0·498
150	−0·866	0·259	0·02	0·966	0·90	−0·88	1·08	−0·865
180	−1	0	0	1	1	−1	1	−1

The close agreement between the second and last columns indicates that the error in writing $\cos\theta$ for $\cos\theta_1$ is a fraction of 1 per cent, and that the true values of θ differ from the approximate values by only a few minutes. Thus the error incurred in substituting θ_1 for θ in (11) is very small, particularly as the term involving θ is multiplied by $2a/3r$, which is about 1/40.

§ 3. APPARATUS

The apparatus is shown in figure 3, and consists of a circular piece of plate glass supported by a vacuum desiccator lid. This communicates, by means of a

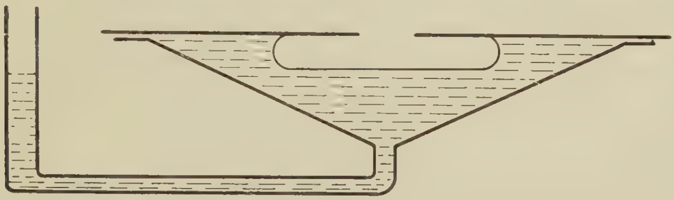


Figure 3.

ground-glass stopper, with a glass tube 1 inch in internal diameter. The whole is supported on a levelling table. A hole 2 inches in diameter passes through the centre of the glass plate.

§ 4. PROCEDURE

The apparatus, after removal from a bath of concentrated sulphuric acid, is thoroughly rinsed in hot boiled water and then dried by means of an electric hair drier. After placing on the levelling table it is filled with the liquid under test, care being taken to avoid touching the apparatus. Some of the liquid is drawn off through the side tube, which serves as a manometer, the large diameter minimizing surface-tension effects. This withdrawal diminishes the level in the circular hole and develops a bubble, which indicates the need of any further levelling of the apparatus. Continued removal of the liquid causes the manometer level to drop and the bubble to grow until the latter is about 5 cm. in diameter, when the level remains practically constant, whereas the bubble continues to grow. This indicates that the base of the bubble has become plane, for no change which could be detected by the unaided eye occurred in the depth while the bubble grew to a diameter of 12 cm. The manometer level is then read by a travelling microscope, for it indicates the level of the base of the bubble. Liquid is next added to the manometer to raise the level above the plate, and thereby to develop the drop. This is allowed to grow to the same diameter as the bubble, and the height in the manometer is read again.

The next problem is the determination of the levels of the surfaces of the glass plate so as to calculate the height of the drop and depth of the bubble. At first it seemed likely that the liquid could be adjusted until the level of the underside of the plate coincided with the surface, when the manometer could be read, but

this was impracticable. Direct measurement is out of the question, as, by the nature of the problem, the capillary rise in the manometer must be accounted for. Finally, the levels were found by placing a spherometer, with its central leg projecting and locked by a nut, astride the hole in the plate. The plate is 0·71 cm. thick, and the spherometer leg was found to project 0·835 cm. below the lower face.

The liquid levels were adjusted by adding the liquid, a drop at a time, to the side tube until contact between spherometer and the surface occurred. A section, of the surface is then as in figure 4, which indicates that the angle of contact will

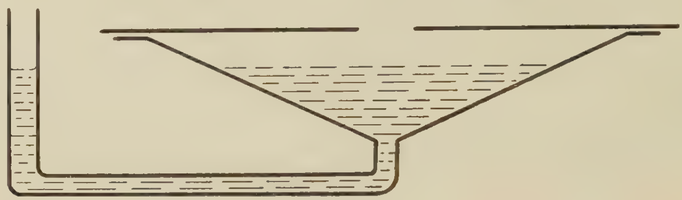


Figure 4.

have a negligible effect on the surface. The level of the manometer is then read, 0·835 cm. is added to give the manometric height corresponding to the lower face of the plate, and a further 0·71 cm. added gives the manometric height of the upper face, from which h_1 and d_1 can be deduced.

(*Note*: If the actual height of the plates is required in the manometer, it can be easily obtained by providing the plate with a 1″ hole, and allowing the liquid to ascend this a little way. As the manometer is 1″ in diameter, the levels will be alike, and the exact values are easily found.)

§ 5. RESULTS

The values for each liquid were obtained at intervals of one week. Benzene was used on three distinct occasions, and gave the same values for height of drop and also for depth of bubble, but the values of each did not give the actual surface tension. This led to the introduction of the spherometer. The aniline and benzene were of A.R. standard, the water was tap-water.

	T	θ	Temperature (° c.)
Aniline	46·7	36	9
Benzene	27·4	$11\frac{1}{2}$	19
Water	73·13	$9\frac{1}{2}$	19

Discrepancies between these results and those generally accepted may be due to

(*a*) Contact being with plate glass.

(*b*) The bubble surface being in contact with air saturated with vapour, whereas this may not be the case with the drop.

(*c*) The liquids, except water, not having been freshly prepared. Cos $11\frac{1}{2}°$ is 0·98, which means an error of about 0·5 dyne in taking θ as 0 for benzene, a quantity which may be overlooked. For water, the difference between $T \cos 9\frac{1}{2}°$ and $T \cos 0$ is about 1 dyne. These results are to be regarded as preliminary, further experimental work having been suspended owing to the prevailing conditions.

§ 6. CONCLUSION

The method has the following advantages:

(*a*) It is statical.

(*b*) The apparatus is easily cleaned.

The development of the bubble in these experiments is really a particular case of the Jaeger bubble experiment, and shows that in the latter the bubble is only to be regarded as spherical for tubes of very small diameter.

The writer hopes to publish more results later, both with the apparatus as described and also with the top of the glass plate covered by a desiccator lid, which will enable the apparatus, including the manometer, to be evacuated.

Such results will give the surface tension and angle of contact of a liquid in contact with its own vapour.

REFERENCES

Burdon, 1940. *Surface Tension and the Spreading of Liquids* (Cambridge University Press), p. 12.

Ferguson, 1913. *Phil. Mag.* **25**, 509.

OPTICAL REFRACTION PATTERNS, PART I: THEORY

By R. S. RIVLIN, M.A.

(Communication from the Staff of the Research Laboratories of The General Electric Company, Limited, England)

Received 3 *April* 1941

ABSTRACT. In an earlier note the writer has described the formation of patterns by refraction of light, from a point source, at the surface of a ground surface of quartz parallel to the (0001) plane and has indicated that the method has been applied to the study of ground and scratched surfaces of quartz and other materials (Rivlin, 1940). In the present paper a mathematical theory of the formation of these refraction patterns is given. Formulae are deduced and curves are given showing the relation between the intensity distribution in the refraction pattern formed by a surface uniformly covered by small facets, having any specified distribution of orientations, and this distribution

of orientations. From a measurement of the light-intensity distribution in the refraction pattern, the law of facet distribution with orientation may be calculated. It is intended, in subsequent papers, to discuss the application of these results to the interpretation of various refraction patterns.

§ 1. INTRODUCTION

The reflection of light by rough surfaces has been the subject of extensive theoretical study by Bouguer (1760), Berry (1923), Pokrowski (1924), Schulz (1925) and Bloch (1939), on the assumption that these surfaces consist of large numbers of small refracting facets. The first four authors made some assumption about the distribution law for facets of various orientations, while the last-mentioned author has deduced from experimental observations the equivalent facet distributions obtaining in the special case of reflection by a road surface.

The " reflectogram " method, in which the intensity distribution in the light reflected from a surface is used to obtain some information about the orientation of various parts of the surface, has been used by a number of workers; the method is essentially an extension of the optical goniometer.

We shall here investigate the intensity distribution in the light which is refracted out of a rough plane surface when a point source of light is viewed through this surface. A distribution of light intensity is then seen, due to the formation of a large number of virtual images of the point source by the various facets of which the matt surface may be considered to consist. This luminous pattern will be called the *refraction pattern* of the surface. Refraction patterns of ground and scratched surfaces of a number of crystalline materials have been studied by the present writer (Rivlin, 1940), and the results obtained were of sufficient interest to warrant a mathematical analysis of the formation of these refraction patterns.

It is assumed here that the surface roughness can be considered to consist of a large number of small plane facets of various orientations, each facet being sufficiently large, however, for the diffraction effects to be neglected. We first consider, in § 2, the refraction pattern obtained in the hypothetical case when the surface is covered by facets all having the same orientation. This case has been discussed, to some extent, by de Gramont (1935), but it is included here as it forms the basis of the subsequent work. It appears that for this case the refraction pattern consists of a single spot of light whose position depends on the orientation of the facets.

In § 3 we discuss how for any given facet orientation this position can be altered. In § 4 the distribution of light intensity for any given distribution of facet orientations is found, and it is shown how the distribution law of facet orientations can be calculated from the observed light-intensity distribution.

It is proposed in subsequent papers to describe the refraction patterns for certain surfaces, and to discuss the inferences which can be drawn therefrom.

§2. REFRACTION BY A SINGLE SET OF SMALL PLANE PARALLEL FACETS

Supposing we have a number of small plane parallel facets standing on the surface AB of a slab of material of refractive index μ, which has two plane parallel bounding surfaces AB, CD (see figure 1). The surface AB is viewed by an eye

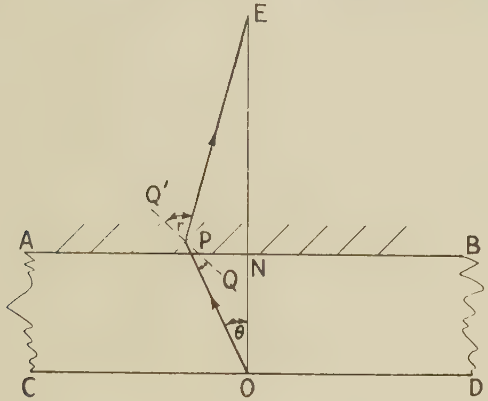


Figure 1.

situated at E and illuminated by a point source of light O on the lower surface CD of the slab. We shall take EO normal to the surfaces AB, CD of the slab. Figure 1 is a plane section through EO and parallel to the line of greatest slope of the facets. Light will be refracted to E from a point P of the surface only if the following two conditions are satisfied :—

(i) the incident ray PO, the normal PQ to the facet at the point P and the line PE (which would then be the direction of the refracted ray), all lie in the same plane;

(ii) the relation

$$\sin i/\sin r = 1/\mu$$

holds, where i and r denote the angles OPQ, EPQ′ respectively.

The first condition limits the position of the point P on the surface AB of the slab to the section of surface by the plane of figure 1, while the second condition limits the point P to a unique position.

With the notation :—

l = thickness of the slab (= ON).

x = distance of eye from the surface of the slab (= EN),

α = inclination of normal to facet to the direction of ON,

PN is given by the equations

$$\sin i/\sin r = 1/\mu, \qquad \ldots\ldots(1)$$

$$PN = x \tan (r-\alpha) = l \tan (\alpha - i). \qquad \ldots\ldots(2)$$

In the limiting case when the eye is at infinity we obtain $r = \alpha$. The value of i, and hence of PN/l, can then be easily found. In figure 2 the value of PN/l for values of α from 0 to 90° is plotted for various values of μ.

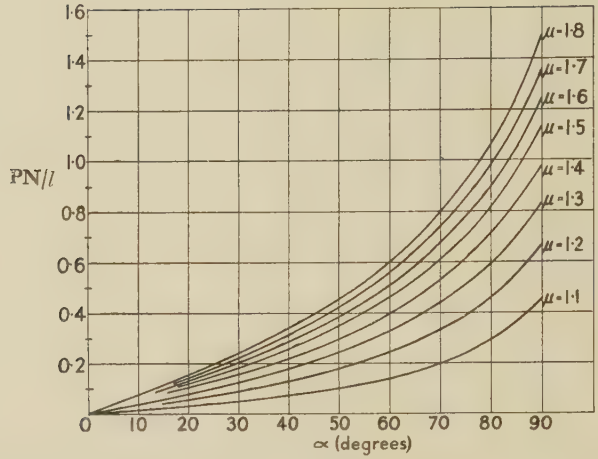


Figure 2. Relation between position on refraction pattern and facet inclination.

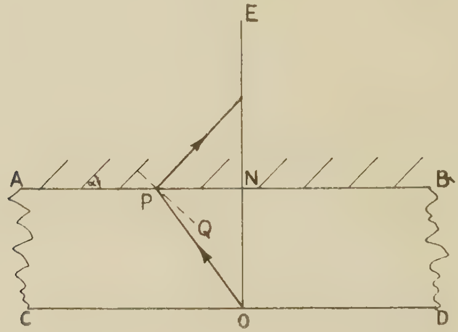


Figure 3.

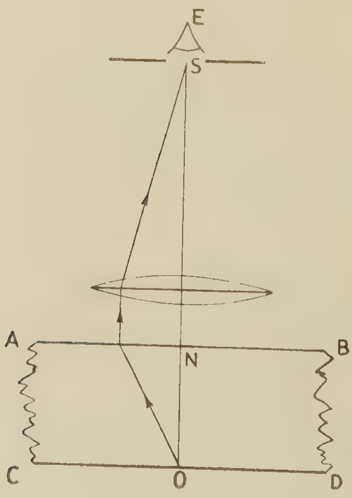


Figure 4.

The nearest position of the eye to the surface at which an image of O can be seen by refraction at a facet of inclination α is given by the construction of figure 3. The eye is then in such a position that it receives a ray leaving the facet in a glancing angle. We see from this that the eye is in a position to see the light refracted from facets of all inclinations up to 90° only if it is at infinity. In practice a close approximation is obtained if the eye is at a distance away from the surface much greater than the thickness of the slab. Alternatively, the arrangement of figure 4 may be used. A lens is placed over the slab and the pattern is viewed through a stop S placed at the focus of the lens. This restricts the light reaching the eye to those rays which leave the surface of the slab in a direction normal to the surface.

§ 3. METHODS OF ALTERING THE DIMENSIONS OF THE REFRACTION PATTERN

From equation (2) it appears that the dimensions of the refraction pattern are proportional to the thickness of the slab, and they may therefore be varied by varying the thickness of the slab. It is, however, not always convenient to do this, and we shall therefore discuss other methods for varying the dimensions of the refraction patterns.

It is evident that if PN, for any particular facet slope, is greater than AN, the corresponding refracted ray will not appear. This may be avoided by covering the surface of the slab with a layer of liquid of suitable refractive index.

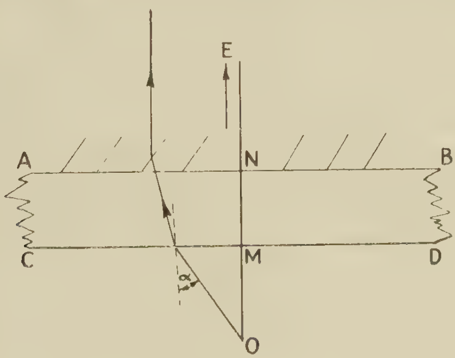


Figure 5.

By suitably choosing the refractive index of the liquid, the dimensions of the refraction pattern may be reduced to any given extent. In the limiting case, when the refractive index of the liquid is equal to that of the material under investigation, the refraction pattern will reduce to a point at N.

If the surface is covered with a liquid of refractive index $\mu_1 (<\mu)$, it is apparent that in the special case when the emergent ray is parallel to ON, PN/l is given as a function of α by substituting μ/μ_1 for μ in equation (1). The corresponding values of PN/l and α are thus given for various values of μ/μ_1 by the curves of figure 2. In the case when $\mu_1 > \mu$, the same argument applies, but the refraction pattern is reversed about the centre N.

Again, the dimensions of the refraction pattern can be increased by taking the point source some distance below the lower surface CD of the slab under investigation. The lower surface should then be polished or covered with a liquid of refractive index equal to that of the material of the slab, to avoid scattering or irregular refraction at this surface. The arrangement is shown in figure 5. It can easily be seen that the position of P corresponding to any given value of α is moved outwards (from N) from its position given by the analysis of § 2 by a distance $a \tan \gamma$, where a is the distance of the source O from the base of the slab and $\sin \gamma = \mu \sin (\alpha - i)$.

§ 4. THE FORMATION OF REFRACTION PATTERNS OF SURFACES ON WHICH STAND FACETS OF VARIOUS INCLINATIONS

We have seen that the refraction pattern of a set of plane parallel facets is a single point whose position depends on the orientation of the facets. Now, if the surface studied is uniformly covered by facets having a variety of orientations (i.e., the fraction of any small area on the surface, covered by facets in any specified range of orientations, is independent of the position of the small area on the surface), a light pattern will be produced which is determined by the distribution of facet orientations.

Let us suppose that the surface is viewed normally through a small square aperture which forms with the point N a pyramid of semi-vertical angle ϵ. Now we shall calculate the intensity of light reaching the aperture from an elementary area dS of the surface situated about a point P. As shown in figure 1, the position of the point P may be defined by spherical polar co-ordinates (R, θ, ϕ), where R is the distance of P from O, θ is the angle PON and ϕ is the angle between the projection of OP on to the plane CD and an arbitrary direction OX in this plane. The orientation of any facet may be defined by two angles α and β. α is the angle of inclination to ON of the normal to the facet. β is the angle between the projection of this normal on to the plane XOC and OX. dS is assumed to be so small that it subtends, at any point of the aperture, an angle which is small compared with ϵ. Then light from O, falling on the area dS, will be refracted into the pyramid formed by P and the aperture only from facets whose normals lie in a small solid angle.

Now, assuming that the source radiates light uniformly in all directions, the amount of light falling on dS from the source at O is proportional to

$$dS \cos \theta / R^2,$$

which for a slab of given thickness is proportional to

$$dS \,.\, \cos^3 \theta.$$

Suppose the surface is uniformly covered with facets, the distribution law of their orientations may be expressed as follows:—

The fraction of the area dS on which stand facets whose orientations lie in the

range

$$(\alpha, \beta), \ (\alpha, \beta+\Delta\beta), \ (\alpha+\Delta\alpha, \beta), \ (\alpha+\Delta\alpha, \beta+\Delta\beta)$$

is given by the formula $f(\alpha, \beta) . \Delta\alpha . \Delta\beta \sin\alpha$. We note that $\Delta\alpha . \Delta\beta . \sin\alpha$ is the solid angle traced out by the normals.

Hence, neglecting the partial reflection which takes place when light is refracted at a facet, the amount of light refracted out of the area dS into the pyramid of semi-vertical angle ϵ is proportional to

$$dS . \cos^3\theta . f(\alpha, \beta)\Delta\alpha . \Delta\beta \sin\alpha, \qquad \ldots\ldots(3)$$

where $\Delta\alpha$, $\Delta\beta \sin\alpha$ are the vertical angles of the " pyramid of normals to facets " corresponding to the pyramid of refracted light of semi-vertical angle ϵ.

We shall now derive the relation between $\Delta\alpha$, $\Delta\beta$ and 2ϵ.

In figure 6, we see how the angle of refraction varies for a small variation $\Delta\alpha$ of the inclination of the facet normal to ON, the incident ray remaining, of course, unchanged.

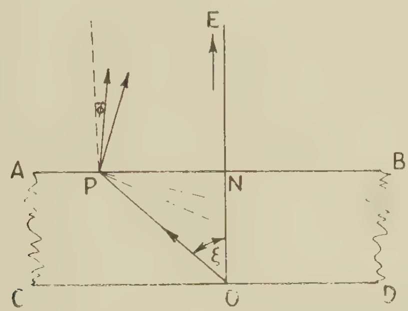


Figure 6.

Let ξ, ψ be the angles of inclination of the incident and refracted rays respectively to ON. Then the relation between the angle of incidence and the angle of refraction is given by

$$\sin(\alpha-\xi)/\sin(\alpha+\psi)=1/\mu,$$

whence
$$\Delta\alpha=\frac{\Delta\psi . \dfrac{1}{\mu} . \cos(\alpha+\psi)}{\cos(\alpha-\xi)-\dfrac{1}{\mu}\cos(\alpha+\psi)}. \qquad \ldots\ldots(4)$$

Now, taking the point from which the surface is viewed far from the surface, we have $\alpha+\psi=\alpha$ and $\Delta\psi=2\epsilon$, which gives, with equation (4),

$$\Delta\alpha=2\epsilon\Big/\left(\mu\frac{\cos i}{\cos\alpha}-1\right),$$

where i is substituted for $(\alpha-\xi)$.

To obtain the relation between $\Delta\beta$ and ϵ it may be noted that the refracted ray, incident ray and normal to the facet must be coplanar, and, since the incident ray remains fixed, must be given by rotating figure 6 about OP through a small angle.

If figure 6 is rotated about OP through the small angle η, the refracted ray

and facet normals each trace out a sector whose vertical angles in planes perpendicular to the paper and passing through the refracted ray and facet normals respectively are given by resolving the rotation in directions in the plane of the paper perpendicular to those of the refracted ray and facet normal respectively.

Therefore the vertical angle of the pyramid of refracted rays is given by $\eta \sin (r-i) = 2\epsilon$ and the vertical angle of the pyramid of facet normals is given by

$$\Delta\beta \sin \alpha = \eta \sin i.$$

Hence

$$\Delta\beta \sin \alpha = \frac{2\epsilon \sin i}{\sin (r-i)} = \frac{2\epsilon \sin i}{\sin (\alpha - i)},$$

since the eye is at infinity, so that $r = \alpha$.

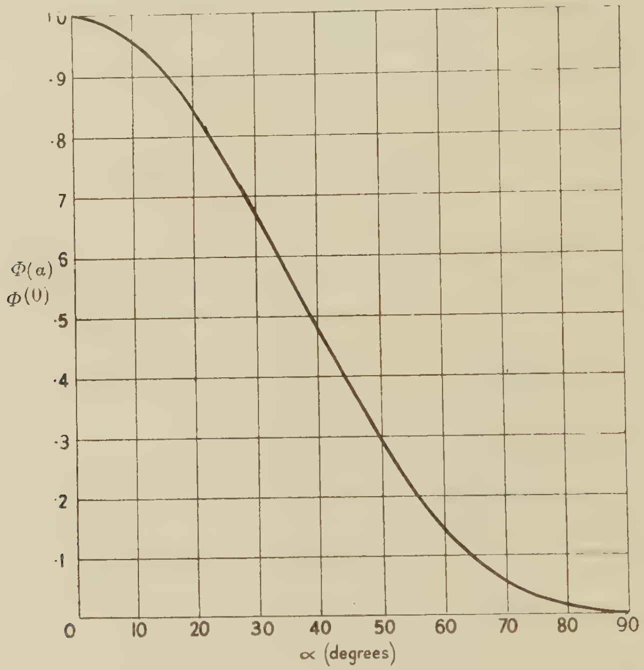


Figure 7. Variation of $\frac{\Phi(a)}{\Phi(0)}$ with a for values of μ between 1·1 and 1·8.

Thus, from equation (3), we have that the light refracted out of the area dS and into the viewing aperture is proportional to

$$\frac{dS \,.\, \cos^3 \theta \,.\, f(\alpha, \beta) 4\epsilon^2 \sin i}{\sin (\alpha - i)\left(\frac{\mu \cos i}{\cos \alpha} - 1\right)}.$$

This formula must be modified by a factor R_i, which is the fraction of the incident light refracted through the surface of refractive index $1/\mu$ when the angle of incidence is i.

For a polished surface this can be obtained from the formulae of Fresnel, which give

$$R_i = 1 - \tfrac{1}{2}\left\{\frac{\sin^2(i-r)}{\sin^2(i+r)} + \frac{\tan^2(i-r)}{\tan^2(i+r)}\right\}$$

for unpolarized light.

Thus the distribution of light over the surface when it is viewed through a small distant aperture is proportional to $\Phi(\alpha) \,.\, f(\alpha, \beta)$, where, since $\theta = \alpha - i$,

$$\Phi(\alpha) = \frac{\cos^3(\alpha - i)\sin i}{\sin(\alpha - i)\left(\dfrac{\mu\cos i}{\cos\alpha} - 1\right)} \,.\, R_i. \qquad \ldots..(5)$$

i is, of course, related to α by the equation

$$\sin i/\sin\alpha = 1/\mu.$$

Now, for a slab of thickness l, α is related to PN/l by the equations (1) and (2), and hence, by using these equations in combination with equation (5), we can evaluate the intensity distribution for any given facet distribution, or *vice versa*. Thus, in figure 7, $\Phi(\alpha)/\Phi(0)$ is plotted against α. The value of the Fresnel reflection factor was obtained from the numerical tables of Moon (1940). The curve given is approximately correct for all values of μ between 1·1 and 1·8. Using this curve in conjunction with the curves of figure 2, $\Phi(\alpha)/\Phi(0)$ may be plotted against PN/l for any value of μ in the range 1·1 to 1·8.

The form of the function $f(\alpha, \beta)$, and hence the equivalent distribution-law of facet orientations, can be deduced by comparing the observed intensity distribution with the distributions given by the curves of figures 2 and 7.

REFERENCES

Berry, E. M., 1923. *J. Opt. Soc. Amer.* **7**, 627.
Bloch, A., 1939. *Trans. Illum. Engng Soc.* **4**, 113.
Bouguer, 1760. *Essai d'Optique* (Paris).
Fresnel, 1812. *Ann. Chim. (Phys.), Paris*, **17**, 194 and 312.
de Gramont, A., 1935. *Recherches sur le Quartz Piezoélectrique* (Editions de la Revue d'Optique Théorique et Expérimentale).
Moon, P., 1940. *J. Math. Phys., Boston*, **19**, 1.
Pokrowski, G. I., 1924. *Z. Phys.* **30**, 66.
Rivlin, R. S., 1940. *Nature, Lond.*, **146**, 806.
Schulz, H., 1925. *Z. Phys.* **31**, 496.

A NEW TREATMENT OF THE THEORY OF DIMENSIONS

By G. BURNISTON BROWN, M.Sc., Ph.D.,
University College, London *
* Now at University College of North Wales, Bangor

Received 24 *May* 1941

ABSTRACT.† A new treatment of the theory of dimensions is proposed, based on consideration of the manner in which numbers are introduced as symbols for certain aspects of nature. It is shown that there are only two direct ways of doing this, the measurement of length and the measurement of time. These are not independent, as the first involves the second ; they are related by a universal constant which is here called the *constant of interaction*, in preference to the " velocity of light ". This leads to a very simple formulation of the dimensions of physical quantities in terms of symbols representing length and time measurements only, and to a new use for dimensional equations in checking theoretical predictions as to the relations between different physical quantities such as mass and energy.

§ 1. INTRODUCTION

JEAN BAPTISTE JOSEPH FOURIER introduced the theory of dimensions into physics in his famous *Théorie analytique de la chaleur* of 1822. His reasons for so doing are stated (Fourier, 1878) as follows :—

" It must now be remarked that every undetermined magnitude or constant has one *dimension* proper to itself, and that the terms of one and the same equation could not be compared, if they had not the same *exponent of dimensions*. We have introduced this consideration into the theory of heat in order to make our definitions more exact, and to serve to verify the analysis ; it is derived from primary notions on quantities ; for which reason, in geometry and mechanics, it is the equivalent of the fundamental lemmas which the Greeks have left us without proof."

Since Fourier's time the subject has been considerably developed. Sir A. W. Rücker (1888) maintained that all physical quantities could be expressed in terms of five units—length, mass, time, magnetic permeability and temperature : Tolman (1917) has suggested length, mass, time, electric charge and entropy : the principles of dimensional homogeneity have been expounded by Buckingham (1914). There is further considerable literature on the subject, but it is true to

† This paper contains the substance of some lectures given at University College during 1939.

say of it all, as Fourier said originally—" It is derived from primary notions .... left us without proof."

Now it is on just such notions which underly explicitly or otherwise the development of physical science that those interested in the philosophy of science must concentrate critical attention.

The treatment developed in this paper is the result of critical consideration of the acts of measurement whereby numbers are introduced into nature, and follows the application of the same method to the problems of electric and magnetic induction (Brown, 1940 a and b). It appears that there are only two fundamental measurements, those of length and time, and this has led to the view that all physical quantities must be capable of expression in terms of length and time alone. It was interesting, therefore, to find that Maxwell (1881) mentions this possibility, although he does not seem to have favoured its adoption. Lord Kelvin (1889) also considered the possibility, saying: " There is something exceedingly interesting in seeing that we can practically found a metrical system on a unit of length and a unit of time. There is nothing new in it, since it has been known from the time of Newton, but it is still a subject full of fresh interest. The very thought of such a thing is full of many lessons in science that have scarcely yet been realised, especially as to the ultimate properties of matter."

The treatment adopted in this paper had as its original stimulus two definitions by Sir Arthur Eddington. The first was his emphasis on the fact that " the whole subject matter of exact science consists of pointer readings and similar indications " (Eddington, 1928), which is really a definition of *exact* science, and the second was his definition of the electric charge $\boldsymbol{e}$, viz., that $\boldsymbol{e}$ is only manifested in the presence of other charges, and in particular, for simplicity, in the presence of *one* other charge, and it is a measure of the mutual interaction.

Consideration of the first led to the view that the only measurements that physicists actually require to make are either of space-like or time-like intervals, and consequently that dimensional reasoning should be in terms of only L and T, and the second (since Eddington's definition applies equally to M, the mass) to the view that the dimensions of charge and mass should be the same, since the methods of measurements of their respective interactions are identical in theory, e.g. by a torsion balance. The new treatment of dimensional theory which follows is more in line with the present philosophy of physics than the classical " intuitive " presentation.

§2. POINTER READINGS

In exact science, numbers are introduced into nature by certain conventional and historic processes. If we exclude the process of pure counting, by which, for instance, we find the number of Jupiter's moons, these conventional processes are found to consist of visual observations of the alignment of two marks, together with counting. In *exact* science these are the only observations that are necessary, and this has led Eddington to describe a physicist as a man who requires no

sense organs beyond one colour-blind eye, and to call the objects of his emasculated attention *pointer readings*.

The physicist claims that he measures many things—mass, force, potentials and so on—but what he actually observes are pointer readings, and this, together with the faculty of counting, is all that is required for that part, and it is important to remember that it is only a part, of science called *exact* science. A hint in this direction might have been obtained by recalling that mass is a coefficient invented by Newton, force is the hypothetical cause of change, and potential is a pure Western-European myth. None of them could, therefore, be measured directly.

Now consideration will show that there are only two direct ways of making use of pointer readings in the process of introducing numbers into nature: in one case the alignment is of two different pairs of marks; in the other case one of the marks at least, and usually both marks, must be the same in each pair of alignments. The first type of observation leads to the measurement of a length or space interval and the second to a measurement of a time interval. In the case of measurement of mass by a balance, or resistance by a bridge method, although a pointer reading is observed, it is not this reading which introduces a number, i.e. it is not by counting marks on this scale that the number is obtained. The number is got by pure counting elsewhere.

Consider first the simple measurement of length. This involves the use of a material scale which is placed in contact with both marks on the object to be measured. The zero mark on the scale is aligned with one of the marks on the object, and then the mark on the scale which is in contact with the second mark on the object is noted. By a process of counting marks on the scale between the two marks with which alignment has been made, a number is arrived at which is defined to be the length between the two marks on the object. Usually the observer moves from one point to the other to observe the two alignments, so that a time interval is necessarily involved in the measurement of a length. To avoid movement the observer may remain at rest at the point of zero alignment and arrange mirrors in such a way as to superimpose an image of the distant alignment upon the zero position, so that both may be observed with the same glance. But even in this case, as we know, a time interval is involved. This is usually expressed by saying that "light takes time to travel". In this paper it will be most important to avoid this popular form of speech. We shall emphasize that, in the case of light, nothing observable has ever been shown to travel at all. We shall not, therefore, make a hypothesis for which there is no evidence whatever; we shall restrict ourselves to the discussion of interaction, which is what we observe.

Now the question which arises in connection with the above discussion of measurement is this: is interaction instantaneous? In other words, if a disturbance is produced at the point occupied by the observer, does its interaction with a distant object occur at the same time (on the observer's clock) as that at which the disturbance was caused? It is, of course, well known that this is not the case. The experiment that shows this is one of the Fizeau-wheel type. A disturbance

is made at the position of the slotted disc ("light passes through a slot") which interacts with an object at a known distance ("illuminates a mirror"), and this interaction itself causes another interaction at the disc ("the light is reflected back to the disc"). It is easily shown that the original disturbance and the second interaction (at the disc) are not instantaneous, and by measuring the time interval and assuming that interaction is independent of direction in space, we infer that the interaction at the mirror, distant l, say, from the disc, took place at a time l/c later than the original disturbance measured on the observer's clock. If care is taken that interaction with other objects is avoided ("free space"), then c turns out to be a universal constant which will here be called the *constant of interaction*. In centimetres and second units the value of c is found to be approximately 3×10^{10}.

Considering next the measurement of time: how do we put numbers into the fleeting time of consciousness? By a method, equally a matter of convention as that by which we measure length. In each case we cannot introduce numbers until we have defined what are to be regarded as *equal* intervals. In the case of length we assume that the scale does not change in length as it is moved about or when it has different orientations in space, and in this way we can mark off equal intervals of distance. In the measurement of time we can define equal intervals as those given by a pendulum swinging in a vacuum in certain specified conditions. The present method of measuring meridian transits of heavenly bodies and then *correcting* them from considerations of Newtonian theory which itself rests on experiment, and therefore on some presumably better definition of equal time intervals, seems unsatisfactory (Brown, 1935). In any case, what is observed is the intermittent succession of alignments between the same pair of marks, one of which is, or has been, in motion. The times between successive alignments of the marks are arbitrarily defined to be equal, even if they differ from the impressions of consciousness. The observation of alignments between the same pair of marks or between pairs of which one is always the same (hour-glass, water clocks, etc.), can be made at the position of the observer without moving or involving interaction with objects at a distance: the observation of the alignment between different pairs, which is necessary for the measurement of length, involves interaction at a distance, and consequently a time interval.

We conclude, therefore, that exact science concerns the introduction of numbers into nature by certain conventional operations together with the process of counting. What are actually observed are pointer readings: when two successive observations consist in noting the alignment of different pairs of marks, a length is said to have been measured; when the successive observations consist in noting the alignment of the same pair of marks, a time is said to have been measured. The observation of different pairs of marks involves movement of the observer, or, if he remain stationary, interaction with the distant point of alignment. The minimum time associated with the measurement of a length l by

a stationary observer is l/c, where c is the constant of interaction. A little consideration will show that this remains true even if the length is calculated by trigonometry from a much shorter base line. The successive observation of the same pair of marks does not require the observer to move, nor does it require any interaction with objects at a distance, so that a time measurement does not involve a length. A "pure" time measurement is therefore possible; a pure length measurement is not. Owing to the fact that c is found to be a universal constant, every measurement of length, even in the ideal case—a theoretical experiment—involves a time interval of l/c. If we now make the following definition:

Equivalent measurements are measurements which in similar physical circumstances result in the same number,

we can say that the measurement of a length l is equivalent to a measurement of a time interval l/c multiplied by the constant of interaction.

At this stage it will be convenient to make some more definitions:—

Primary measuring operations are the operations described above for introducing numbers as symbols of spatial and temporal intervals, given in, or inferred from, consciousness.

Secondary measuring operations are all other operations of measurement, such as weighing or reading a thermometer.

Measurement: the measuring operations, together with the substitution of the numbers so obtained in the formula defining a physical quantity, constitute the measurement of that quantity.

A physical quantity is anything that can be measured.

Magnitude of a physical quantity is a number obtained as the result of measurement.

§ 3. DEFINITION OF DIMENSIONS

Let the operation of measurement of a length be symbolized by L and the operation of measurement of a time be symbolized by T. If the results of all measurement can be calculated from numbers obtained by one or both of these two operations, we must be able to express the measurement of any physical quantity in terms of L and T. The measuring operation involves, of course, a unit, and in what follows the c.g.s. system will be assumed for convenience; but the treatment is independent of such an assumption.

Taking first the example of an area measurement, we obtain the number representing the area by multiplying the numbers obtained by two different length measurements. We can represent this by $L^1 \times L^1$ or L^2. In the same way a volume measurement is represented by L^3. An angular measurement requires two measurements of length, but the resulting numbers are divided, and so this is represented by L^0. Similarly a velocity is LT^{-1} and an acceleration LT^{-2}. With this convention, the index numbers indicate the way in which the numbers resulting from operations of measurement are treated in arriving at the magnitude of the physical quantity in question: they do *not* necessarily represent the number

of separate measurements of length and time which are required for determining this magnitude.

We can now define what is meant by the dimensions of a physical quantity :—

Dimensions: the dimensions of a physical quantity consist of one or both of the symbols for the two primary operations of measurement of length and time (L and T), together with the respective indices representing the powers to which the numbers so obtained are raised in order to conform to the definition of the quantity.

Ratio of physical quantities: the ratio of two physical quantities is obtained by dividing their respective magnitudes and dimensions.

Thus the ratio of two lengths will be a pure number. It is important to notice that although the number obtained by a length measurement is the ratio of the length of the object to that of the standard, this does not make all lengths pure numbers (dimensionless). The length of an object in *exact* science is not something given directly in consciousness. It is a number obtained by a conventional process called measurement. Now the length of the standard is not got by measurement but by *definition*, and so it is correctly represented dimensionally solely by unity. The length of a rod measured by it is, say, 10 . L. On forming the ratio we have, therefore, that the ratio of the length of the rod to that of the standard is 10 . L.

Recalling now the argument of § 2, which led to the view that although a pure time measurement is possible, a pure length measurement is not, we see that in measuring a length we cannot avoid, in fact, measuring a space-time interval. Thus we can always arrive at the number representing it in two ways, either by a length measurement or by measuring the corresponding time interval and multiplying by $\boldsymbol{c}$, the constant of interaction. Consequently we can write

$$\mathrm{L} = \boldsymbol{c}\mathrm{T}$$

and look upon it as the fundamental dimensional relation, to be interpreted as follows: the symbol for a time measurement equivalent to a length measurement is obtained by dividing the symbol representing the length measurement by the constant of interaction $\boldsymbol{c}$. This means that in dimensional formulae we are entitled to substitute $\boldsymbol{c}\mathrm{T}$ for L.

§ 4. DIMENSIONS OF MECHANICAL QUANTITIES

As soon as we leave kinematics and turn to other branches of physics, we deal with quantities which are not so directly measurable, e.g. force, electric charge, etc., and which are defined by equations containing quantities which can be measured. Such equations may be called *defining equations*. The general procedure is to use these equations to define the unit and the dimensions. In cases where a physical quantity is determined by an equation expressing a proportionality (e.g. force proportional to rate of change of momentum), the

unit of the quantity is defined so as to make the numerical value of the constant of proportionality unity. In an exactly similar manner the dimensions of the quantity are chosen so as to make the constant of proportionality dimensionless. It can then, and only then, be dropped from the equation.

Now Newtonian mechanics is founded on the inference, from experiment, that force is required only to change velocity and not to maintain it. A convention for measuring force and mass was adopted by Newton which, since it is used as a defining equation, is now written

$$(\text{force}) = (\text{mass}) \times (\text{acceleration}).$$

This, of course, is not sufficient, since there are two unknowns and only one equation. Newton, however, added his law of gravitation, giving, for the case of equal masses,

$$(\text{force}) \text{ is proportional to } \frac{(\text{mass})^2}{(\text{distance})^2}.$$

We have now two equations and two unknowns, and so we can find either the force or the mass. Writing the above in the form

$$(\text{force}) = GM^2/r^2,$$

where G is some constant, and substituting for the force from the first equation, we get

$$GM/r^2 = \text{acceleration}.$$

Using square brackets to denote " dimensions of ", and taking G to be a dimensionless constant, we have

$$\begin{aligned}[\mathrm{M}] &= [\text{acceleration}]\,[\text{length}]^2 \\ &= \mathrm{L}^3\mathrm{T}^{-2}.\end{aligned}$$

Later, in considering electricity and magnetism, it will be shown that we are correct in taking G to be dimensionless, since it is necessary for consistency, and in agreement with the definition of the ratio of physical quantities.

In consequence of the above dimensional formula we have:

mass	$\mathrm{L}^3\mathrm{T}^{-2}$	action	$\mathrm{L}^5\mathrm{T}^{-3}$
force	$\mathrm{L}^4\mathrm{T}^{-4}$	density	$\mathrm{L}^0\mathrm{T}^{-2}$
energy	$\mathrm{L}^5\mathrm{T}^{-4}$	angular momentum	$\mathrm{L}^3\mathrm{T}^{-3}$,
momentum	$\mathrm{L}^4\mathrm{T}^{-3}$		

and so on.

At first sight it appears strange that, as Lord Kelvin said, the fourth power of a linear velocity is the proper measure of a force, and " the square of an angular velocity is the proper measure of density ". Kelvin, however, worked out ingenious illustrations: " find the velocity with which a particle of matter must be projected, to revolve in a circle round an equal particle fixed at such a distance from it as to attract it with a force equal to the given force. The fourth power

of this velocity is the number which measures the force ". And again, with regard to density, Kelvin showed that if a satellite rotates very close to the surface of a spherical mass, then

$$(\text{density of sphere}) = 3/4\pi\,(\text{satellite's angular velocity})^2.$$

Thus reassured, we may pass on to consider electricity and magnetism.

§ 5. DIMENSIONS OF ELECTRIC AND MAGNETIC QUANTITIES

It has already been noticed in § 1 that Eddington's emphasis on the fact that e, the electric charge, is a measure of the interaction of two charges leads directly to the view that M, the mass, and m, the magnetic pole-strength, are also measures of interaction, and as the measurement by which they are evaluated could be exactly the same in each case (e.g. using a torsion balance), and as the formula is also the same (inverse-square law), the dimensions for each should also be identical. This is clearly shown by writing

$$[\text{force}] = \left[G\frac{M_1 M_2}{r^2}\right] = \left[\frac{e_1 e_2}{K r^2}\right] = \left[\frac{m_1 m_2}{\mu r^2}\right].$$

Now K and μ are dimensionless numbers (Brown, 1940 a), for in each case they are defined as the ratio of numbers obtained by similar measurements, performed first with the material in question and then with a vacuum. The dimensions which represent the measurement are therefore similar, and when the ratio is derived they cancel and we are left with pure numbers. It cannot be too strongly emphasized, in view of the nature of past literature, that K and μ are what we define them to be, and not in some mysterious way indicative of the " physical nature " of electricity and magnetism (Tolman, 1917).

K and μ being dimensionless, the consistency of the above equations is satisfied if G is taken to be dimensionless, and it should be noted that its appearance in the equation is merely due to the fact that an arbitrary standard of mass already existed which was not defined in a similar manner to the unit of charge and pole. For consistency we should have taken unit mass to be such that if two unit masses were concentrated each at one of two points in a vacuum far removed from other bodies, and separated by the unit of distance, the force of attraction would be the unit of force.*

If this consistent procedure had been followed, the constant G would not have appeared, and all three equations would have been isomorphous for a vacuum (i.e. when all other masses, charges, or poles are far removed from the two in question). It is clear, therefore, that we are correct in taking G to be dimensionless, and, in fact, it is easily shown that $\sqrt{(1/G)}$ is the ratio of this unit of mass to the arbitrary gram ($\sqrt{1/G} = 3{\cdot}9 \times 10^3$)—a dimensionless ratio of numbers obtained by similar measuring operations (in this case counting).

* It would, of course, have been more consistent to base the whole of physics on the gravitational unit of mass ($1{\cdot}5 \times 10^7$ gm.). This is defined in terms of length and time only.

The dimensions of the electric charge and magnetic pole are therefore the same as those of mass, viz. L^3T^{-2}. We easily derive the following dimensions :—

Quantity	Dimensions: Electrostatic system	Dimensions: Electromagnetic system
Electric charge	L^3T^{-2}	L^2T^{-1}
Electric intensity and induction	LT^{-2}	L^2T^{-3}
Electric potential and e.m.f.	L^2T^{-2}	L^3T^{-3}
Electric current	L^3T^{-3}	L^2T^{-2}
Electric resistance	$L^{-1}T$	LT^{-1}
Electric capacity	L	$L^{-1}T^2$
Electric inductance	$L^{-1}T^2$	L
Magnetic pole	L^2T^{-1}	L^3T^{-2}
Magnetic field and induction	L^2T^{-3}	LT^{-2}
Magnetic flux	L^4T^{-3}	L^3T^{-2}
Intensity of magnetization	L^0T^{-1}	LT^{-2}

This table illustrates the great simplicity of the system proposed, which includes the absence of fractional indices.

§ 6. DIMENSIONS OF HEAT QUANTITIES

The subject of heat differs from other branches of physics in that it was long thought that no connection existed between mechanics and heat. Consequently mechanics was not put into heat at the start in the same way that it was put into electricity and magnetism, i.e. heat was not measured by means of the changes in force that it produces. This has made the determination of the dimensions of heat quantities less obvious than in the corresponding cases of electricity and magnetism, since no defining equations existed connecting the units of heat and temperature with mechanical quantities. The dynamical theory of heat, however, provides us with two equations, the equivalence of heat and energy, and Kelvin's absolute dynamical scale of temperature (the perfect gas equation). The entropy-probability relation and Stefan's law do not express entropy or temperature in terms of any mechanical measurement, and are not, therefore, suitable as defining equations.

Now according to the dynamical theory, heat is not merely proportional to energy, but is itself a form of energy, and consequently a measurement which evaluates energy can be used equally to evaluate heat. The dimensions must therefore be the same, so that any factor used for conversion from arbitrary heat units to energy units must be a dimensionless number. Thus we can write

$$[\text{heat}] = L^5T^{-4},$$

so that we have [mass] [specific heat] [temperature] $= L^5T^{-4}$.

Now in the classical treatment, specific heat is dimensionless because, by definition, it is the ratio of the quantity of heat required to raise the temperature of a body by a given amount to the quantity of heat required to raise the temperature of an equal mass of water by the same amount. Thus, whatever symbols we

use to denote the measurement of a quantity of heat, when we form the ratio the symbols cancel and we are left with a pure number. We have, therefore,

$$L^3T^{-2}\,[\text{temperature}] = L^5T^{-4}$$

or

$$[\text{temperature}] = L^2T^{-2}.$$

The dimensions of temperature are thus those of a square of a velocity. This agrees, as it should, with the kinetic theory, for we can interpret this result in a similar manner to the way in which Lord Kelvin interpreted the result that force is represented by the fourth power of a velocity (§ 4). We use the kinetic theory and imagine an ideal gas as the thermometric substance. It can then be shown that the temperature is proportional to the mean-square velocity of the molecules. In particular, two different materials are at the same temperature if, when they are placed in thermal contact with the thermometric substance, the mean-square velocity of its molecules is the same in each case.

Another way of showing this might be as follows:—Consider two masses M_1 and M_2 of fluid, having specific heats s_1 and s_2, and let their temperatures be t_1 and t_2. Now let them interact by placing them in contact. Let t_3 be the temperature after the interaction and s_3 the specific heat. In the ideal case, with no heat loss or chemical action, we write for the quantities of heat

$$s_1M_1t_1 + s_2M_2t_2 = s_3(M_1 + M_2)t_3.$$

Now let the masses interact in another way. Let their respective velocities be v_1 and v_2, and let them collide, remaining together after the collision and having a common velocity v_3. In the ideal case, where no heat is produced, we write for the quantities of energy

$$\tfrac{1}{2}M_1v_1^2 + \tfrac{1}{2}M_2v_2^2 = \tfrac{1}{2}(M_1 + M_2)v_3^2.$$

In the first case we have a heat equation; in the second, a mechanical equation. If we are to treat heat as a branch of mechanics, then it is clear from the similarity in form of the two equations, and the fact that specific heat is dimensionless, that

$$[\mathrm{t}] = [\mathrm{v}]^2$$

i.e. temperature must have the dimensions of a velocity squared.

Following from these we find

$$[\text{thermal conductivity}] = L^4T^{-5}$$
$$[\text{entropy}] = L^3T^{-2}.$$

The dimensions obtained above for temperature using the classical definition of specific heat are consistent with the modern definition of specific heat as energy per gram per degree. They can also be obtained by using the ideal gas equation (corresponding to Kelvin's absolute dynamical scale) as a defining equation, provided that it is correctly expressed with reference to a specified quantity M of material, i.e.

$$pv = \boldsymbol{k}MT,$$

where the symbols have the usual meaning and $\boldsymbol{k}$ is a constant.

It may be valuable at this point to discuss the two kinds of dimensionless

quantities. Firstly, there are those that are *defined* as ratios of the results of similar measurements. These, of course, do not alter in value with change of units. They include specific gravity, specific heat, specific inductive capacity (dielectric constant), and permeability. Then there are those that can be shown to be dimensionless factors like the gravitational constant and the mechanical equivalent. These arise if the measurement of a physical quantity is made using two different systems of units, only one of which is directly related to the units of length and time: in the case of G, the use of the arbitrary gram as well as a gravitational unit of mass (which is implied as soon as length and time units are specified): in the case of J, the use of heat units and mechanical units for energy. When *both* systems are separately connected with the units of length and time, as in the case of the E.S. and E.M. systems, the constant factor is a power of c and is not dimensionless.

These latter dimensionless quantities are dimensionless *on the dynamical theory*. For example, when everything is measured in mechanical units, J is the ratio of two energies. When, however, we do not measure everything in mechanical units (and this is the usual case), then these quantities vary with the units adopted.

§ 7. THE RATIO OF THE UNITS OF THE ELECTROSTATIC TO THE ELECTROMAGNETIC SYSTEM

The E.S. and the E.M. systems are united by theory. The theory rests on the hypothesis that an electric current is physically equivalent to a uniformly moving stream of electrostatic charges. This hypothesis has, of course, received adequate confirmation from direct experiment. This enables us to define a physical quantity on the two systems, i.e. to specify two different methods of measurement and to assert that they are measures of the *same physical quantity*. Thus in the definition of equivalent measurements (§ 2), we require theory to help us to decide what should be "similar physical circumstances". Now, since measurements are symbolized by dimensions, equivalent measurements should be symbolized by equivalent dimensional formulae. But we have seen that all measurements can be represented by L and T raised to various powers, and further, we have found the time measurement equivalent to a length measurement, given by

$$\mathrm{L} = c\mathrm{T}.$$

We are, therefore, in a position to examine the equivalence, or otherwise, of different measurements.

Let us take the convenient case of the measurement of the capacity of a condenser. This, as we know, can be performed merely by the use of a measuring scale on the E.S. system; on the E.M. system a clock is required in addition. Thus this particular example suggests immediately that there must be some equivalence between a length and a time measurement if they are to result finally in the same number.

On the E.S. system the measurement is represented by L, and on the E.M. system it is represented by $L^{-1}T^2$.

Now we can obtain this latter formula by multiplying the former by L^2/L^2 and substituting the equivalent time measurements,

$$L = L \times \frac{L^2}{L^2} = \boldsymbol{c}^2 L^{-1} T^2.$$

This shows that the E.S. measurement (L) is equivalent to the E.M. measurement ($L^{-1}T^2$), provided that this measurement includes the multiplier $\boldsymbol{c}^2$. Thus, if we multiply the number derived from E.M. measurement by $\boldsymbol{c}^2$, we shall get the same number as we should derive from E.S. measurement. Let these numbers be $C_{\text{E.M.}}$ and $C_{\text{E.S.}}$ respectively, then, if the experiment is made, we shall expect to find

$$C_{\text{E.M.}}\boldsymbol{c}^2 = C_{\text{E.S.}}$$

or

$$\sqrt{\left(\frac{C_{\text{E.S.}}}{C_{\text{E.M.}}}\right)} = \boldsymbol{c} = 3 \times 10^{10} \text{ cm./sec.}$$

As is well known, this is fully borne out by experiment, and provides, therefore, evidence for the correctness of the theory which led to it, viz., that a stream of uniformly moving electric charges is physically equivalent to (indistinguishable by any measurement from) an electric current.

§ 8. A DARK-ROOM EXPERIMENT

Let us imagine the determination of the ratio of the capacity of a condenser on the E.S. and E.M. systems to be performed in a dark room. The galvanometer must be a pointer instrument arranged so that the position of the pointer on the scale could be felt, and all scale divisions deeply engraved so that they could be counted by touch. The clock might be a metronome. As the experiment is only a rough one, the observer might be allowed to ignore the logarithmic decrement. After a period of practice, the observer produces the result $(3 \times 10^{10} \text{ cm./sec.})^2$.

Now, whatever may have happened in the dark room during the interval necessary for the production of the result, it seems clear that, as all light has been carefully excluded, and as no measurement of any velocity of any radiation has been made, it would require considerable temerity to assert that what has been measured is the velocity of light. From the point of view here adopted, it would seem much more likely to represent the " rate of exchange " between two sets of conventional operations for putting numbers into nature.

§ 9. THE MASS-ENERGY AND MASS-MOMENTUM RELATIONSHIPS

Continuing with the viewpoint that

$$L = \boldsymbol{c}T$$

is an equation expressing the fundamental dimensional equivalence, we can pursue further the question of equivalent physical measurements. Consider the dimensions of energy,

$$[\text{energy}] = L^5 T^{-4}.$$

By substituting from the fundamental equation we can write

$$[\text{energy}] = \boldsymbol{c}^2 \mathrm{L}^3 \mathrm{T}^{-2}$$

or

$$[\text{energy}] = \boldsymbol{c}^2\,[\text{mass}].$$

Thus an energy measurement and a mass measurement are equivalent if the number found in the latter case is multiplied by $\boldsymbol{c}^2$. Or, in other words, if in some physical interaction we find that the mass decreases and the energy increases, this is the relation which the numbers produced by the respective measurements should bear to one another. This relation has, of course, been predicted theoretically and verified experimentally.

In a similar way we get

$$[\text{momentum}] = \boldsymbol{c}\,[\text{mass}]$$

with a corresponding interpretation.

Since mass and electric charge have the same dimensions, it is clear that we also have the equation

$$[\text{energy}] = \boldsymbol{c}^2\,[\text{electric charge}].$$

Thus if an interaction should be observed in which a measurement of electric charge is possible before the interaction, and afterwards less charge or no charge manifests itself, we should expect to find an increase in energy equal to the number representing the decrease in charge multiplied by $\boldsymbol{c}^2$. Or equally, if the charge-loss and energy-gain did not balance we might expect an increase in mass. It is important, however, not to fall into the old error of thinking that the study of dimensions can perform what is really the province of that creative thought which produces what we call physical *theories*. Dimensional analysis in the past has been used to check equations of physical terms: if the treatment in this paper is valid, dimensional analysis also allows us to check physical theories. For if it is not possible to transform the dimensions of one physical quantity into the dimensions of another physical quantity by means of the fundamental equation of equivalence $\mathrm{L} = \boldsymbol{c}\mathrm{T}$, then the respective measurements cannot be equivalent, and consequently the numbers produced will not be the same. Thus the theory which leads to the view that one of the physical quantities can " turn into " the other cannot receive verification. Mass may " turn into " energy but not into angular momentum or action.

In this connection it is of interest to note that all physical quantities to which conservation laws apply in ideal cases are measured by equivalent processes:

Quantity	Dimensions	Equivalent dimensions
Energy	$\mathrm{L}^5\mathrm{T}^{-4}$	$\boldsymbol{c}^4\mathrm{L}$
Momentum	$\mathrm{L}^4\mathrm{T}^{-3}$	$\boldsymbol{c}^3\mathrm{L}$
Mass, charge (E.S.) and entropy	$\mathrm{L}^3\mathrm{T}^{-2}$	$\boldsymbol{c}^2\mathrm{L}$
Charge (E.M.)	$\mathrm{L}^2\mathrm{T}^{-1}$	$\boldsymbol{c}\mathrm{L}$

§ 10. CONCLUSION

The treatment of the theory of dimensions advocated above results in a great simplification in the usual expressions for the dimensions of physical quantities. Instead of five different symbols, only two are required, L and T, and these represent the only *direct* measuring operations which are available to us for introducing numbers into nature by means of pointer readings. Fractional indices disappear and the indices required do not exceed six, including zero. Another use for dimensional analysis follows from the new approach, for it can be used as a check upon theory where it concerns the possibility of one physical quantity "turning into" another, and the equivalence of measures of the same physical quantity on different systems.

The importance of calling $\boldsymbol{c}$ the constant of interaction has been stressed. The word velocity is better reserved for all cases in which something observable moves. In the case of light and "electromagnetic radiations" generally, nothing has ever been observed moving. Thus it may be better to regard light, not as something which travels through space from the luminous object to the eye, either particle or wave, but as that part of the continual interaction across space which affects the organ of sight. In the case of other continual interactions across space, such as that of the moon and the tides, we have learned to do without the unnecessary hypothesis of something travelling through space. The reason why the "velocity of light" plays such an important part in modern physics is that it is not the velocity of anything. The "waves of light" are a purely conceptual device for explaining, predicting and calculating various effects of interaction. As atomic particles are never observed, but only inferred from interaction, it should not cause any surprise that a wave theory which has proved useful for describing one form of interaction should prove useful for describing other forms. Thus light waves and mass waves are only conceptual devices for dealing with interaction, which is what is measured.

The fact that *exact* science deals with pointer readings, and that these may be reduced to those of length and time only, explains how it is that the world of exact physics can be reduced to nothing but the properties of a *space-time* continuum. Everything in nature that cannot be symbolized by a number introduced by these two arbitrary conventions is *left out*.

§ 11. ACKNOWLEDGEMENT

The author is greatly indebted to Sir Arthur Eddington for criticism which caused him to treat the section on heat in greater detail.

REFERENCES

Brown, G. B., 1935. "The Limits of Science." *Sci. Progr.* **116**, 729.

Brown, G. B., 1940 a. "A New Treatment of Electric and Magnetic Induction." *Proc. Phys. Soc.* **52**, 577.

Brown, G. B., 1940 b. "Fundamentals of Classical Electric and Magnetic Theory." *Nature, Lond.*, **145**, 789.

Buckingham, E., 1914. *Phys. Rev.* **4**, 345.

Eddington, Sir A. S., 1928. *The Nature of the Physical World*, p. 252.
Fourier, J. B. J., 1878. *The Analytical Theory of Heat* (Cambridge University Press).
Kelvin, Lord, 1889. *Popular Lectures and Addresses*, **1**, 103.
Maxwell, J. C., 1881. *Electricity and Magnetism*, **1**, 4.
Rucker, Sir A. W., 1888. *Proc. Phys. Soc.* **10**, 37.
Tolman, R. C., 1917 a. *Phys. Rev.* **9**, 237.
Tolman, R. C., 1917 b. *Phys. Rev.* **9**, 242.

THE DIMENSIONS OF PHYSICAL QUANTITIES

By W. E. DUNCANSON, M.Sc., Ph.D.,
University College, London*
* Now at University College of North Wales, Bangor

Received 1 *November* 1940

ABSTRACT. As a basis for the discussion of the dimensions of physical quantities, the three usual quantities (length, time and mass) are taken as indefinables and the mechanical quantities are defined in terms of these. The dimensions follow directly from the definition of a quantity and indicate the type of measurements that have to be made to determine the magnitude of those quantities. It is shown that a further indefinable is necessary for defining electrical and magnetic quantities, and reasons are given for choosing electric charge for this purpose. It is pointed out that when the dielectric constant and the permeability are equated to unity for a vacuum it amounts to using these two quantities as indefinables, which, although in some cases convenient from practical considerations, is not necessary.

For defining thermodynamical quantities, only the three indefinables, mass, length and time, are used. The introduction of Boltzmann's constant, as a numerical scale factor, is not necessary, but gives a more convenient unit of temperature. The significance of the introduction of such scale factors is also discussed from a more general point of view.

Finally, it is shown that three indefinables are not necessary for measuring the magnitude of mechanical, and hence thermodynamical quantities, and that these can be quite adequately defined in terms of length and time only, while electrical and magnetic quantities are measured in terms of length, time and charge.

§ 1. INTRODUCTION

The subject of this paper has been discussed by a large number of writers both in letters and longer articles and, in referring to the various opinions that have been expressed, it will not be convenient to give all the references but, at the outset, to mention three classical contributions by Rücker (1888/90), Maxwell (1904) and Tolman (1917) as having considerably assisted in clarifying the subject. However, many difficulties still remain, especially in the subject of electricity and magnetism. It seems that the difficulty and disagreement over the subject of dimensions are partly due to the lack of a clear statement of the meaning of the terms used, and of the assumptions that have been made. Therefore a brief account will be given now of the assumptions that are to be made in this paper.

Firstly, the axioms and simple results of mathematics will be assumed. These are discussed by Russell (1903) and Tolman (1917), and we shall not retrace the ground here. Two terms " magnitude " and " physical quantity " are particularly important for our discussion. Following Russell and Tolman, " magnitude " is taken as undefined, it being assumed there is general agreement on the meaning conveyed by such a term, and " quantities are to be regarded as magnitudes which have been particularized by the specification of spatial and temporal conditions " (Tolman, 1917). Later we shall have to discuss the meaning of further terms such as the " dimensions " and " nature " of a physical quantity, but we shall proceed now to consider the meaning of the measurement of the magnitude of physical quantities and the processes involved. To make such measurements it is necessary to have certain undefined quantities in terms of which other quantities can be measured. The quantities that can be taken as indefinables are arbitrary but, in practice, certain quantities are found to be more suitable than others

Tolman has dealt with this problem at some length and has given reasons for the suitability and sufficiency of the three indefinables, mass, length and time, for the measurement of the mechanical properties of matter. When we turn to other branches of physics, and as physical knowledge increases, it is possible that phenomena may be encountered which will require further indefinables for their complete specification. On the other hand, it is possible, and in many ways desirable, that as our knowledge increases the integration of physics will become more complete and a smaller number of indefinables will suffice, and it may be hoped that ultimately only one such indefinable will be necessary.

To begin with, then, let us take mass, length and time as our indefinables and consider whether it is necessary in our present state of knowledge to add further indefinables for the complete expression of physical phenomena; towards the end of the paper the possibility of reducing the number will be considered.

§2. INDEFINABLE QUANTITIES AND METHOD OF MEASUREMENT

Before being able to determine the magnitude of a physical quantity, a statement has to be made concerning the magnitude of the indefinables, and the conventions to be adopted for making comparisons with them. Such a statement will, in effect, describe the procedure to be adopted for the measurement of a mass, a length, or a time.

Standard of length. Two points A and B are marked on a plane surface. A straight line is drawn between A and B and the line is taken as our unit of length. Let us call this unit of length the metre; that is, the line has a length of 1 metre, or the distance between the two points is 1 metre. The convention for comparing other lengths with this standard is as follows:—If a body, whose length is to be compared with the standard length, is placed so that one end coincides with the mark A on the standard, and it is found that the other end coincides with mark B, then the body is said to have a length equal to that of the standard, that is, it has

a length of 1 metre. It is to be noted that this is the length in a certain direction in the body, namely, the direction of the straight line joining the two points in the body which are coincident respectively with A and B. By further application of the method of coincidences, lengths of 2, 3, 4, . . . metres can be measured. Also, dividing the standard into a number of equal lengths, by a similar method, lengths can be measured to a fraction of the standard.

Standard of time. The unit of time can be taken as a complete rotation of the earth on its axis relative to the fixed stars, that is, a sidereal day—and can then, with the help of the pendulum, be subdivided into a smaller unit, the second. Times are then measured by counting beats of a pendulum or by some equivalent method.

Standard of mass. A certain piece of material is taken and this is said to have unit mass, which we shall call the kilogramme. The convention for measuring the mass of a body is not so straightforward as that for measuring length or time. Most conventions make direct or indirect use of the concept of force. But if we are to adhere to three indefinables, then force can be defined only after laying down the conventions for measuring mass, length and time. In theory a simple convention would be as follows. If a body produces an acceleration on a second body n times that produced by the standard when placed the same distance away from the second body, it is said to have a mass n times that of the standard. In practice such a measurement is difficult and the procedure has been to make certain assumptions, such as Newton's laws of motion, which involve the concept of a force and then to describe a method for the measurement of mass (Filon, 1926–7, and 1938; Pendse, 1937, 1939 and 1940).

§ 3. SIGNIFICANCE OF DIMENSIONS *

We are now in a position to measure other physical quantities in terms of these three indefinables. The type of measurement that we make for the determination of the magnitude of a particular quantity will be decided by the definition of that quantity. For example, speed is defined as the rate of change of position. To measure the speed of a body then involves two experimental operations, namely, a comparison of a length with the standard length and a comparison of a time with the standard time, and the mathematical operation of division. Again, to measure a momentum we have to measure a mass, a length and a time. The results of the mass and length measurements are multiplied together and the product is divided by the result of the time measurement.

The symbols [L], [T], [M] will be used to represent the operations of comparison with the standard of length, time and mass respectively. Mathematical operations will be represented by the usual symbols. Thus the operations involved in measuring the magnitude of the speed of a body are represented by $[L]/[T]$ or $[LT^{-1}]$.†

* See Brown (1940) for a discussion of this point.

† This notation does not differentiate between scalars and vectors, velocity also being represented by $[LT^{-1}]$; but it will be assumed that when the term " dimensions of a quantity " is used in the strict sense it will include information concerning the scalar or vector nature of that quantity.

Similarly, the operations for measuring the magnitude of the momentum are represented by $[MLT^{-1}]$, and so on for all the quantities associated with the mechanical properties of matter.

Such a group of symbols is said to represent the dimensions of the quantity measured; or, in other words, the dimensions of a quantity denote the type of measurements that have to be made in order to determine the magnitude of the particular quantity concerned. It is often assumed that the dimensions indicate the physical nature of a quantity. If this is so, then it means that the physical nature of a quantity is determined by the type of measurements made on it, where measurement is meant to include not only the operations of comparison with the indefinables, but also any necessary mathematical operations indicated by the dimensional expression. Or, to stress again the close relationship between the definition and the dimensions of a quantity, the nature of a quantity is already specified by the definition of it. Hence the phrases *nature of mass*, *nature of length*, *nature of time* have no physical meaning as these quantities have not been defined, and the nature of a quantity which is defined is determined by that definition; for example, *the nature of momentum* means nothing more than is contained in the definition, stated in simple terms, mass times velocity. Therefore if there is doubt concerning the dimensions of a quantity this means, not that there is some hidden mystery as to the complete nature of that quantity, but merely that the quantity concerned has not been unambiguously defined. This will become more clear in the following discussion of electrical and magnetic quantities.

§ 4. ELECTRICAL AND MAGNETIC QUANTITIES *

Let us now consider whether electrical and magnetic phenomena require further indefinables for their specification. We shall not delay here to describe all the elementary experiments which give meaning to such terms as *electrified body* and *magnetized body*, but shall proceed to express in mathematical terms the results of these experiments and to define a number of suitable quantities.

Firstly, we shall consider the results of the experiment which shows that there is a force of attraction, or repulsion, between two electrified bodies. It will be assumed that the two bodies are small and are at a distance large compared with their size. Experiment shows that the force between two electrified bodies is inversely proportional to the square of the distance between them, that is,

$$F \propto \frac{1}{r^2},$$

or

$$F = \frac{\alpha}{r^2},$$

where α is independent of r. In all practical cases this force is so much greater than the gravitational force that the latter may be neglected.

Further information can be gained about α. It is found to be independent

* Cf. a series of letters (*Nature, Lond.*, **139**, 45, 473, 676 and 844), and a recent Physical Society discussion (Brown, 1940).

of the masses of the two bodies, but dependent on the medium in which the bodies are embedded and on the electrification of the two bodies. This latter phrase will now be given more precise statement.

Consider three small electrified bodies A, B and C. B and C are separately placed at a distance *d*, large compared with their size, from A and the separate forces measured; then B and C are placed together at the same distance *d* from A. It will then be found that the force on A is equal to the sum or difference of the two forces exerted separately by B and C—the sum when B and C repel one another, and the difference when they attract. Again it can be shown, in a similar manner, that the force on A is dependent also on the electrification of A. Hence the expression for the force between two bodies A and B can be more definitely expressed as

$$F=\frac{q_A q_B}{\kappa r^2}, \qquad \ldots\ldots(1)$$

where q_A is a quantity, called the charge, associated with body A, q_B is a similar quantity associated with body B and κ is a quantity, called the dielectric constant, associated with the medium.

It must be noted that q_A, q_B, and κ are symbols introduced to represent certain properties of the bodies and the medium, but as yet not given precise definition. As q_A and q_B are the same type of quantity, then in equation (1) there are two new quantities q and κ which require to be defined. Until they are defined their dimensions cannot be stated. Here a difficulty arises. There are two quantities to be defined, but only one equation for the purpose. Before considering the implications of this let us, by way of introduction, consider two methods that have been adopted by many physicists for dealing with such quantities. Firstly, there are those who state that κ is a pure number (Brown, 1940), which for convenience is equated to unity for a vacuum, and then the product of two equal charges is defined as the product of the force between them *in vacuo* and the square of their distance apart,

i.e., $$q^2=[Fr^2]=[ML^3T^{-2}],$$

$\therefore$ $$q=[M^{1/2}L^{3/2}T^{-1}].$$

The appearance of non-integral powers in the dimensions is disconcerting as they have no physical meaning. It may be suggested that M, L, and T are not convenient quantities in terms of which to express electrical quantities, but it will be shown later that there is a more fundamental reason for the difficulty.

A second approach to the problem is to admit that κ has dimensions which at the present cannot be determined, and q is then expressed as

$$q=[M^{1/2}L^{3/2}T^{-1}\kappa^{1/2}],$$

the suggestion being that κ may have dimensions such that the non-integral powers of M and L would disappear (Starling, 1929).

It will now be shown that, in effect, these two approaches amount to very much the same thing. In the first case κ is conveniently equated to unity for

a vacuum which, as Rücker (1888–90) pointed out, makes it easy to forget its dimensions, and then the square of the charge is put numerically equal to the value of the force multiplied by the value of the square of the distance. Furthermore, equating κ to unity for a vacuum does not define κ,* nor does the more general statement that κ is a pure number. In physics any quantity which is a pure number represents either the process of counting or the ratio of two quantities which have the same dimensions. One might define κ for a medium as the ratio of the force between two charges *in vacuo* to the force between them in the medium, but, in effect, this only defines the ratio of the dielectric constant for the medium to that for a vacuum.

In the second method there is ambiguity concerning the dimensions of both κ and q, which means that they have not been defined unambiguously. In fact, neither can be measured until their dimensions are known, as it is not known what type of measurements to make. This, at first sight, seems unreasonable, for it may be contended that q can be measured, but in actual fact it is only the product, $q\kappa^{-1/2}$, which is defined—or rather, as will be seen later, the product $q^2\kappa^{-1}$—and this has the dimensions $[M^{1/2}L^{3/2}T^{-1}]$, which amounts to much the same as the first approach. Rücker (1888–90) suggested that κ could be treated as an indefinable,* but in all these cases the problem of the non-integral powers of the dimensions remains.

Tolman (1917) has pointed out that the non-integral powers would disappear if the charge q were regarded as an indefinable but gave no reasons for introducing a fourth indefinable other than those given by Rücker, which do not strike at the root of the problem, and we shall here take the problem a step further.

§ 5. INTRODUCTION OF CHARGE AS A FOURTH INDEFINABLE

The equation for the force between two electrified bodies contains two quantities q and κ which require definition. With only one equation available it is not possible to define both. Moreover, there is no relation known at the present time which involves only one electrical or magnetic quantity; this means, then, that in order to define electrical quantities, one of them must be taken as an indefinable, that is, other electrical and magnetic quantities will be defined in terms of this fourth indefinable together with M, L and T. The question arises as to the most convenient electrical quantity to take as an indefinable.

Firstly, let us consider the possibility of the dielectric constant κ. q^2—not q, let it be noted—can now be measured as the product of a force, the square of a distance and a dielectric constant, but this does not give us a measure of q. In fact, whether κ be regarded as dimensionless, expressible entirely in terms of M, L and T, or indefinable, it is not possible to measure the magnitude of a single electric charge in terms of M, L and T or of M, L, T and κ. Take as an illustration the following definition of the electrostatic unit of charge. "Unit charge is that charge which repels a similar equal charge with a force of 1 dyne when placed a distance of 1 cm. apart *in vacuo*."

* Cf. a further discussion on this in § 8.

Now this is not what it purports to be, namely, the definition of unit charge, but the definition of the product of two unit charges. In other words, one is defined in terms of the other, and not solely in terms of M, L, T (and κ). Further, all other electrical quantities whose dimensions in terms of M, L, T (and κ) contain non-integral powers are defined in terms of a single charge (or its equivalent) as well as in terms of M, L, T (and κ). For example, potential is defined as work per unit charge. On the other hand, resistance, which does not contain any non-integral powers in its dimensions, is defined in terms of the square of a charge—more explicitly, work per unit charge divided by charge per unit time. Hence, as we cannot measure the magnitude of a single electric charge in terms of M, L, T (and κ) alone, and as certain electrical quantities are defined in terms of a single electric charge, then the indications are that a further indefinable is necessary for the complete determination of electrical (and magnetic) quantities. The foregoing discussion shows that κ is not a suitable quantity for this purpose and indicates charge as the obvious choice.

Standard charge. The standard charge is now to be placed on the same footing as the standard length, time and mass. The natural unit to take would be the charge carried by the proton (or the electron). The convention for comparing a charge with the standard would be as follows: a body would have a charge n times that of the proton if it repelled (or attracted) a third charge with n times that force which the proton exerted on it when placed at the same distance from it *in vacuo*.

§ 6. DEFINITION OF ELECTRICAL QUANTITIES

We are now in a position to define the electrical quantities in terms of M, L, T and Q.

Dielectric constant is defined with the help of equation (1), which may be written

$$\kappa = \frac{q_A q_B}{Fr^2}, \qquad \ldots\ldots(2)$$

or in words, κ is defined in terms of the product of two charges divided by the product of a force and the square of a distance. Its dimensions are

$$\kappa = [M^{-1}L^{-3}T^2Q^2]. \qquad \ldots\ldots(3.1)$$

Examples of the dimensions of other electrical quantities are

$$\text{Current} \quad i = dq/dt = [T^{-1}Q]. \qquad \ldots\ldots(3.2)$$

$$\text{Potential} \quad V = dW/dq = [ML^2T^{-2}Q^{-1}]. \qquad \ldots\ldots(3.3)$$

$$\text{Resistance} \quad R = V/i = [ML^2T^{-1}Q^{-2}]. \qquad \ldots\ldots(3.4)$$

$$\text{Capacity} \quad C = q/V = [M^{-1}L^{-2}T^2Q^2]. \qquad \ldots\ldots(3.5)$$

Care must be taken here to distinguish between the two different quantities which are often loosely both termed the capacity of a condenser. Firstly, there is the quantity, which we shall represent by C'—sometimes referred to as the

geometrical capacity of a condenser—whose magnitude is determined by measuring the geometrical dimensions of the condenser and this has the dimensions [L]. Secondly, there is the quantity which we shall denote by C and call the capacity. This is defined by equation (3.5) and it has the dimensions $[M^{-1}L^{-2}T^2Q^2]$. The difference between these two quantities is that C takes into account the presence of the medium and C' does not. Both are useful quantities, but it must be realized that they are different, and confusion of the two will lead to error. C' refers only to the condenser, whereas C refers to the condenser and the medium. The relation between them is given by

$$C=\kappa C'.$$

§ 7. ELECTROMAGNETIC AND MAGNETIC QUANTITIES

We now proceed to define electromagnetic and magnetic quantities. Usually, magnetic quantities are defined in terms of a magnetic pole and then the electric and magnetic quantities are related by means of a constant. It is not necessary to introduce a further indefinable, as these quantities can be defined entirely in terms of M, L, T and Q. It must be noted that no attempt will be made to give the most rigid general definitions, as the aim is to concentrate on the dimensions of the quantities concerned. For this purpose simple definitions will be used deliberately.

Permeability. Let us take two small conductors of lengths ds_a and ds_b placed so that they are perpendicular to the line, of length r, joining them. If currents i_a and i_b respectively are passing through them, then the force between the two conductors is dependent on i_a, i_b, ds_a, ds_b, r and the medium in which they are embedded. For a given medium and constant currents we have that

$$F \propto \frac{ds_a ds_b}{r^2}.$$

As it has not been found possible to introduce a quantity into the equation which will represent the medium and have a definite value for each medium, the following procedure is adopted. It is found that *in vacuo* the force is proportional to the product of i_a and i_b, so that we can write for a vacuum

$$F \propto \frac{i_a i_b ds_a ds_b}{r^2},$$

or

$$F=\mu \frac{i_a i_b ds_a ds_b}{r^2}, \qquad \ldots\ldots(4)$$

where μ is a constant.

Equation (4) is now used for the more general case of any medium, μ being not a constant for a given medium but dependent on i_a and i_b. However, if i_a and i_b are kept constant, μ varies from medium to medium and is termed the permeability of the medium, being defined by equation (4), which can be re-written thus:—

$$\mu=\frac{Fr^2}{i_a i_b ds_a ds_b}.$$

Therefore the dimensions of μ are

$$\mu = [\mathrm{MLQ^{-2}}].$$

Magnetic induction. Consider the force, due either to an electric current or a magnet, on a small conductor of length ds carrying a current di placed at a point P, where di is so small that it does not affect the existing field. It is then found that

$$F \propto di\, ds,$$

or

$$F = B\, di\, ds.$$

If the direction of the current element is such that the force on it is a maximum, then B, thus defined, is called the *magnetic induction* and has the same direction as the force. Its dimensions are

$$B = [\mathrm{MT^{-1}Q^{-1}}].$$

Magnetic field. The magnetic field, H, at a point is defined by the expression

$$H = B/\mu,$$

where B is the magnetic induction at that point embedded in a medium of permeability μ. The dimensions of H are thus

$$H = [\mathrm{L^{-1}T^{-1}Q}].$$

Magnetic moment. We shall take for granted the results of the elementary experiments on the properties of magnets, the location of poles and the definition of the axis of a magnet. Let us consider the effect of a complete plane circuit of small area A carrying a current i_a on a small conductor b of length ds_b carrying a small current di_b at a distance large compared with the size of the circuit. For a given medium it is found that the effect of the circuit on the small conductor is the same as that produced by replacing the circuit by an appropriate magnet of small length with its axis perpendicular to the plane of the circuit. For a given orientation of the circuit and conductor, and for constant di_b and ds_b it is found that the force on b is given by

$$F_i = \beta\mu i_a A, \qquad \ldots\ldots(5)$$

where β is a quantity depending only on di_b and on the geometry and orientation of the system, and is used for convenience rather than inserting the various quantities each time. However, it is found that the force due to the magnet on the conductor is independent of μ and depends only on β and the strength of the magnet. Let us introduce a quantity σ to represent the strength of the magnet so that the force on the conductor is given by

$$F_m = \beta\sigma. \qquad \ldots\ldots(6)$$

Thus σ, which we call the *magnetic moment* of the magnet, is defined by equation (6), where the value and dimensions of β can be obtained from equation (5); or, alternatively, if the magnet is of such a strength as to produce the same force as the circuit on the conductor, then we have

$$\sigma = \mu i A. \qquad \ldots\ldots(7)$$

σ thus has the dimensions

$$\sigma = [\mathrm{ML^3T^{-1}Q^{-1}}].$$

Pole strength. If l is the distance between the two poles of a magnet, then we define a quantity m called the pole-strength of the magnet by

$$m = \sigma / l,$$

the dimensions being

$$m = [\mathrm{ML^2T^{-1}Q^{-1}}].$$

Force between two magnets. To complete our discussion, let us enquire how the force between two magnets depends on μ. The force between a magnet and a small circuit carrying a current i is independent of μ and is given by

$$F = \gamma\sigma i,$$

where γ involves only the geometry and orientation of the system. Now we see from equation (7) and the preceding discussion that the circuit can be replaced by a second magnet with a magnetic moment given by

$$\sigma' = \mu i A;$$

$$\therefore \qquad F = \frac{\gamma\sigma\sigma'}{\mu A},$$

or

$$F \propto \frac{\sigma\sigma'}{\mu}.$$

Thus the force between two magnets is inversely proportional to μ.

In the foregoing we have given one method of defining the various electrical and magnetic quantities. As was emphasized earlier, the definitions are not rigid and omit many important considerations, but our concern has been to set up a consistent and unambiguous set of definitions which will reveal at each point the dimensions of the quantity concerned.

§ 8. SOME REMARKS ON κ AND μ

Before leaving this question of the dimensions of electrical and magnetic quantities there is one further point to which some attention should be paid. From time to time the claim is made that both κ and μ are dimensionless (Brown, 1940). Although we have already dealt with the claim that κ is dimensionless, it is worth while considering this further statement that both κ and μ are dimensionless, as it involves one or two interesting points. Let us consider the two expressions

$$F = \frac{q_A q_B}{\kappa r^2} \qquad \text{......(8)}$$

and

$$F = \mu \frac{i_a i_b ds_a ds_b}{r^2}. \qquad \text{......(9)}$$

If κ and μ are dimensionless, then the right-hand sides of equations (8) and (9) differ in dimensions by the square of a velocity. Therefore it is not possible for both κ and μ to be dimensionless *in the same system*, unless a further quantity with the dimensions of the square of a velocity be introduced in the denominator of equation (8), or in the numerator of equation (9); alternatively a constant c_1

may be inserted in the denominator of equation (8) and a constant c_2 in the numerator of equation (9), where c_1 and c_2 have any dimensions so long as their product has the dimensions of the square of a velocity. The introduction of such constants is artificial and unnecessary. Moreover, those who claim that κ and μ have no dimensions do not really treat them as dimensionless quantities but as *indefinables*. In effect the treatment is as follows:—

Take κ as an indefinable and let it have the value unity for a vacuum. Then the convention for measuring κ for any other material is to measure the ratio of the force between two charges at a given distance apart *in vacuo* to the force between them when placed the same distance apart in the material. Or, alternatively, it may be measured by measuring the ratio of the capacities of a condenser with and without the material between the plates.

In a similar manner μ is taken as an indefinable.

In any case, whatever interpretation is given to the procedure of equating κ and μ to unity for a vacuum, we have shown above that charge is required as an indefinable and then the dimensions of κ and μ become definite.

§ 9. MAGNITUDE OF ELECTRICAL AND MAGNETIC QUANTITIES

So far we have taken the minimum number of indefinables convenient for our purpose and defined the various quantities in terms of these without regard to the magnitude of their units. When we consider this, difficulties and confusion are liable to arise. The units commonly used for various quantities have been chosen, not so much according to a logical scheme as for convenience. As a result more standards are introduced than is necessary. Let us consider, as an example, the method used for measuring the magnitude of a resistance. According to equation (3.4), an operation of the type $[ML^2T^{-1}Q^{-2}]$ has to be performed in order to make this measurement. However, in practice, a system of units is often adopted which takes as a standard unit resistance a certain quantity of mercury in a certain condition, and in order to measure our resistance the only operation necessary is a comparison with the standard. In other words, resistance instead of charge has been taken as an indefinable in this particular case. This is not a very convenient procedure, as the dimensions of a quantity will depend on the particular system of units adopted. In order to overcome this difficulty we will adopt the convention that the measurement of resistance involves not only the comparison with the standard resistance, but also the measurement of this standard in terms of M, L, T, and Q. (This is not a convention as to the meaning of dimensions, let it be noted, but a convention about these secondary standards.) The dimensions of a quantity would then be independent of the particular system of units, and any factors used for converting from one system of units to another would be merely scale factors with no dimensions. For example, there are two other systems of units in common use, namely, the electrostatic and the electromagnetic systems of units, and often there appear in an equation quantities some of which are expressed in e.s.u. and some in e.m.u.

In such a case there also appears in the equation a numerical factor to take account of this. From the foregoing we see that this is merely a numerical factor and has no dimensions. For instance, in Maxwell's equation

$$\frac{d^2E}{dt^2} = \frac{9\times 10^{20}}{\mu\kappa}\nabla^2 E,$$

we have to introduce a factor 9×10^{20} as both μ and κ are given the value unity for a vacuum. In other words, μ is in e.m.u. and κ is in e.s.u. If, however, the same units were adhered to throughout, then the factor 9×10^{20} would disappear. This problem arises again when we come to consider the introduction of Boltzmann's constant into thermodynamical quantities and will be discussed in further detail there.

§ 10. THERMODYNAMICAL QUANTITIES

When we come to thermodynamical quantities we have to consider whether it is necessary to introduce a further indefinable. Rücker was almost convinced that it was not necessary but, because of some doubt, proposed introducing temperature as an indefinable (not his word), and as an indication of his uncertainty he called it a secondary fundamental unit. Tolman accepted Rücker's assumption that a thermodynamic indefinable is necessary but preferred it to be entropy rather than temperature.

Now from the purely physical point of view there need be no limit to the number of indefinables used, but, because the multiplication of the number of such quantities hinders the co-ordination and simplification of physics, we attempt to reduce the number to a minimum. Therefore our aim here is to see whether the subject of thermodynamics requires the introduction of any further indefinables.

Let us examine the problem firstly from simple considerations of the kinetic theory of gases and then from a more general point of view which will take into account the presence of radiation. Simple gas-kinetic calculations give for the pressure on the walls of a vessel containing a gas

$$p = \tfrac{1}{3} n_0 m\overline{c^2},$$

where n_0 is the number of molecules, of mass m, per unit volume and $\overline{c^2}$ is the mean square of the velocities. If n is the total number of molecules in a vessel of volume v then

$$p = \tfrac{1}{3}\frac{n}{v} m\overline{c^2},$$

or

$$pv = \tfrac{2}{3} n \,.\, \tfrac{1}{2} m\overline{c^2},$$

or

$$pv = \tfrac{2}{3} n\overline{E}, \qquad \ldots\ldots(10)$$

where $\overline{E}$ is the mean kinetic energy of the molecules. That is, it is the mean value of the kinetic energy of a large number of molecules, it being impossible to measure the kinetic energy of the individual molecules. $\overline{E}$ is an important quantity in

thermodynamics and is usually designated by a separate symbol and given a separate name. Let us define a quantity θ, called the temperature of the gas, by the equation

$$\theta = \tfrac{2}{3}\bar{E}. \qquad \ldots\ldots(11)$$

The dimensions of θ follow immediately from this definition and are

$$\theta = [\mathrm{ML^2T^{-2}}]\ *$$

that is, it has the dimensions of energy, which is to be expected, as it was made equal to the mean kinetic energy multiplied by the trivial numerical factor 2/3.

Entropy, S, is defined in the usual manner by the equation

$$dS = dE/\theta \qquad \ldots\ldots(12)$$

and is a pure number, as can be seen from equation (11). The dimensions of other thermodynamical quantities are given in the table.

Hitherto we have restricted ourselves by using the kinetic theory as our starting point. Let us turn now to the method of statistics. Without attempting to discuss the subject of statistical mechanics, we shall make use of some of its results on which to base a more general set of definitions. Let us consider an assembly of entities (particles or quanta) with a given total energy. Of the variety of possible distributions in phase space, one distribution will have a maximum probability W. This quantity W is a pure number, being the ratio of the number of different ways in which the " maximum distribution " can be arranged to the total number of possible arrangements of the assembly. A quantity, S, which we shall call the entropy of the assembly, is defined now by the equation

$$S = \log_e W. \qquad \ldots\ldots(13)$$

This again defines entropy as a pure number. Temperature is then defined from equation (12) as the ratio of the energy taken in by a system to its change in entropy. We thus have the same dimensions for thermodynamical quantities as in our previous set of definitions.

§ 11. MAGNITUDE OF THERMODYNAMICAL QUANTITIES

Turn now to a consideration of the magnitude of the units of quantities defined above. This has, so far, been purposely avoided. Our main concern has been to give unambiguous definitions of thermodynamical quantities, leaving till later the introduction of a scale factor to bring the magnitude of the quantities to a more convenient size. Such factors can be divided into two classes:

* Porter in his book *The Method of Dimensions* takes the equation $pv = R\theta$ to refer to unit mass. It can then be written $pv = mR\theta$, where v is the volume of a mass m of the gas. He states that R is a numeric and deduces the dimensions of temperature to be $[\mathrm{L^2T^{-2}}]$. There is no reason for this, especially as R, defined in this way, is a property of the particular gas in question. Even if the equation refers to a gm.-molecule, in which case R is a universal constant and m refers to a gm. molecule of the gas, it must be realized that in actual fact the quantity, gm.-molecule, has no dimensions, but in reality represents a number—namely, a certain number of molecules—as can be seen more clearly from equation (10).

(*a*) A scale factor having a definite numerical value, such as various powers of ten for converting from one unit of length, e.g. the metre, to other units of length such as the centimetre, the kilometre, the angstrom or the XU; or again, a factor of ten to convert from the electromagnetic system of units to the practical system. Such scale factors have a definite value and their numerical value can always be inserted in an equation.

(*b*) The second class of scale factors arises as follows:—Instead of taking a definite numerical factor, a value is taken so as to make the new unit a convenient value for a particular body or substance. This type of scale factor has already been met in § 9 when the magnitudes of electrical and magnetic quantities were being discussed. We now wish to introduce a similar factor to set up a convenient unit of temperature.

According to the definitions of thermodynamical quantities given above, normal temperatures would be of the order of 10^{-14}, so it would be convenient to decrease the magnitude of the unit of temperature by a factor of about this order. Let us introduce a scale factor with a value of 10^{-16} approximately and denote it by $\boldsymbol{k}$. The actual value of $\boldsymbol{k}$ is adjusted so that the difference in temperature between melting ice and steam at the boiling point is 100 units, the new unit being called the degree centigrade. Further, as a matter of convenience, the unit of entropy is increased by a factor $1/\boldsymbol{k}$. Equations (11), (12) and (13) become

$$\boldsymbol{k}\theta = \tfrac{2}{3}\overline{E},$$
$$dS = dE/\theta,$$
$$S = \boldsymbol{k} \log_e W.$$

In accordance with the convention laid down in § 9, the measurement of temperature is meant to involve a comparison with the new unit together with a comparison of the new unit with the original unit of temperature. Thus $\boldsymbol{k}$ is treated as a numeric and has no dimensions.

One possible criticism must be anticipated. It may be argued that in considering electrical quantities care was taken to show that the dimensions of κ and μ could not be suppressed, and yet here, although in the beginning the symbol $\boldsymbol{k}$ has been omitted, it later appears as a scale factor with no dimensions. However, the introduction of κ and μ was necessary as the force between two charges (or currents) varied with the medium and quantities had to be introduced to allow for this change, whereas physically there is no need at all to introduce $\boldsymbol{k}$ into thermodynamics. It is used in order to avoid the trouble of dealing with temperatures of the order of 10^{-14}.

Throughout, it has been tacitly assumed that energy is measured in ergs. Here again a scale factor $1/J$ is introduced, where J has the approximate value $4{\cdot}2 \times 10^7$ and the new unit is called the calorie. The scale factor is adjusted so that the amount of energy required to raise 1 gm. of water 1° C. at 15° C. is equal to one new unit, i.e. to 1 calorie.

The dimensions of the important thermodynamical quantities are collected together in the table below.

Quantity	Dimensions in terms of M, L, T, Q	Dimensions in terms of L, T, Q
Mechanical quantities		
Velocity	$[LT^{-1}]$	$[LT^{-1}]$
Acceleration	$[LT^{-2}]$	$[LT^{-2}]$
Mass	$[M]$	$[L^3T^{-2}]$
Momentum	$[MLT^{-1}]$	$[L^4T^{-3}]$
Force	$[MLT^{-2}]$	$[L^4T^{-4}]$
Energy	$[ML^2T^{-2}]$	$[L^5T^{-4}]$
Electrical and magnetic quantities		
Dielectric constant	$[M^{-1}L^{-3}T^2Q^2]$	$[L^{-6}T^4Q^2]$
Current	$[T^{-1}Q]$	$[T^{-1}Q]$
Potential	$[ML^2T^{-2}Q^{-1}]$	$[L^5T^{-4}Q^{-1}]$
Resistance	$[ML^2T^{-1}Q^{-2}]$	$[L^5T^{-3}Q^{-2}]$
Capacity	$[M^{-1}L^{-2}T^2Q^2]$	$[L^{-5}T^4Q^2]$
Electric field	$[MLT^{-2}Q^{-1}]$	$[L^4T^{-4}Q^{-1}]$
Electric induction	$[L^{-2}Q]$	$[L^{-2}Q]$
Inductance	$[ML^2Q^{-2}]$	$[L^5T^{-2}Q^{-2}]$
Permeability	$[MLQ^{-2}]$	$[L^4T^{-2}Q^{-2}]$
Magnetic induction	$[MT^{-1}Q^{-1}]$	$[L^3T^{-3}Q^{-1}]$
Magnetic field	$[L^{-1}T^{-1}Q]$	$[L^{-1}T^{-1}Q]$
Magnetic moment	$[ML^3T^{-1}Q^{-1}]$	$[L^6T^{-3}Q^{-1}]$
Pole strength	$[ML^2T^{-1}Q^{-1}]$	$[L^5T^{-3}Q^{-1}]$
Thermodynamical quantities		
Temperature	$[ML^2T^{-2}]$	$[L^5T^{-4}]$
Quantity of heat	$[ML^2T^{-2}]$	$[L^5T^{-4}]$
Entropy	$[0]$	$[0]$
Thermal conductivity	$[L^{-1}T^{-1}]$	$[L^{-1}T^{-1}]$
Specific heat	$[M^{-1}]$	$[L^{-3}T^2]$

§ 12. DEFINITION OF MECHANICAL QUANTITIES IN TERMS OF TWO INDEFINABLES ONLY

It was assumed at the beginning of this paper that three indefinables were necessary to describe the mechanical properties of matter. It will now be shown that these properties of matter can be adequately specified in terms of length and time only, and hence electrical and magnetic phenomena in terms of length, time and charge. This suggestion has been made from time to time, but has never gained general approval.

Not all the experiments described in the following have been carried out in practice, but from our knowledge of the properties of matter the results given are those to be expected if the experiments were made.

Consider the following set of experiments:

(*a*) A body is fixed at a point A. A second body B at a distance r from A will

experience an acceleration which will be found to be inversely proportional to the square of its distance from A.

i.e. $$f_B \propto \frac{1}{r^2},$$

or $$f_B = \frac{m}{r^2}, \qquad \ldots\ldots(14)$$

where m is independent of r.

(b) Remove part of A and repeat the experiment. The constant m in equation (14) now changes its value, but it is not altered by removing part of B. Thus the quantity m is associated with some property of A, but not of B, so that we can re-write equation (14) as

$$f_B = \frac{m_A}{r^2}. \qquad \ldots\ldots(15)$$

(c) Split A into two parts A_1 and A_2 and repeat experiment (a). The result obtained would show that

$$m_A = m_{A_1} + m_{A_2}.$$

This quantity m associated with body A we call the *mass* of A. It is defined in terms of f and r from equation (15) and will have the dimensions

$$m = [L^3T^{-2}].$$

The unit thus defined is much too large for practical purposes and it will be convenient again here to introduce a scale factor.

It cannot be said that the quantity defined by equation (15) is the same as the usual quantity called mass. In the one case mass is defined in terms of length and time, whereas in the other it is taken as an indefinable, and, if undefined, then it cannot be said to be the same as something which is defined. Therefore the question of the identical nature of these two quantities does not arise.

The force on a body is now defined as the product of the mass and the acceleration produced on the body, that is,

$$F = mf = [L^4T^{-4}] \qquad \ldots\ldots(16)$$

and so for other mechanical quantities.

The force between two bodies of mass m_A and m_B is then

$$F = \frac{m_A m_B}{r^2}.$$

Objections are raised to this definition of mass on the grounds that mass is of an entirely different nature from length and time (see, for instance, Jeans (1925)). As we have already stated, from the point of view of the physicist the nature of a quantity is determined solely by its definition. Here we have defined mass entirely in terms of length and time.

The suggestion that mass is of a different nature from length and time is

probably due in part to our experience of it, although in actual fact we are never conscious of mass, but of force together perhaps with the area over which it acts, but this does not necessarily mean pressure. However, the important point is that the impression of mass, by the intermediary of a force, is communicated to the body by muscular and other means entirely different from those of time and of length, which are chiefly visual. But the impression conveyed to the body must not be taken as a guide to the physical nature of a quantity. One only has to consider the variety of ways in which the body is sensitive to electromagnetic radiation as the wave-length varies, to see that the body is no reliable guide to the " nature " of a phenomenon.

Another criticism that is raised is that we have two expressions giving respectively the gravitational and the electrical force between two bodies, viz.,

$$F \propto \frac{m_A m_B}{r^2} \qquad \ldots\ldots(17)$$

and

$$F \propto \frac{q_A q_B}{r^2}. \qquad \ldots\ldots(18)$$

In the one case we express mass in terms of length and time only, yet we claim that with a similar relation holding in the electrical case charge cannot be so expressed. It need only be pointed out that, in our definition of mass, equation (17) was not used but is derived only as a result of our definition of mass from equation (15), and of force from equation (16). Equation (15) contains only one mass, whereas no corresponding expression can be found in the electrical case.

§ 13. CONCLUDING REMARKS

The results of this paper lead to the conclusion that the dimensions of a physical quantity depend on the definition of the quantity and the particular quantities which are taken as indefinables. The choice of the latter is, to some extent, arbitrary and there can be no such thing as the absolute dimensions of a quantity, as some writers suggest (Starling, 1929). As we have said earlier, dimensions are merely an indication of the type of measurements that have to be made in order to determine the magnitude of a quantity in terms of the magnitude of certain other specified quantities.

REFERENCES

Brown, 1940. *Proc. Phys. Soc.* **52**, 577.
Filon, 1926–7. *Math. Gaz.* **13**, 146 ; 1938, **22**, 9.
Jeans, 1925. *Mathematical Theory of Electricity and Magnetism* (5th ed.), 14.
Maxwell, 1904. *Electricity and Magnetism* (3rd ed.), **2**, 241.
Pendse, 1937. *Phil. Mag.* **24**, 1012 ; 1939, **27**, 51 ; 1940, **29**, 477.
Rücker, 1888–90. *Proc. Phys. Soc.* **10**, 37.
Russell, 1903. *Principles of Mathematics.*
Starling, 1929. *Electricity and Magnetism* (5th ed.), 385.
Tolman, 1917. *Phys. Rev.* **9**, 239.

GRAVITY METERS

By J. McGARVA BRUCKSHAW, Ph.D, M.Sc.,
Imperial College, London

Lecture delivered 14 *March* 1941

§ 1. INTRODUCTION

The present interest in gravity meters arises from two somewhat widely separated considerations, one academic and the other essentially economic. The former springs from the lack of observations on the variation of gravity over large areas of the earth's surface, particularly over the sea, and has led to efforts being directed to the design of equipment capable of measuring changes of gravity under the difficult conditions existing on board ship, or the somewhat quieter conditions in a submerged submarine. The latter consideration is due to the fact that certain mineral deposits, in particular oil, are associated with large rock structures having a density considerably different from the surrounding rocks. As a result, there is frequently a gravitational anomaly, sometimes attaining values as large as 30 to 40 milligal but, more usually, less than 10 milligal, at the surface above the deposits. The existence of these anomalies has been appreciated for the last 20 to 30 years, and observations of the anomalous gravity values have been made indirectly by means of the Eötvös torsion balance, which measures the magnitude of the horizontal gradient of gravity, and its direction. Obviously, from the gradient at a sufficient number of stations, the gravity changes can be computed by integration. Although this method has met with considerable success, particularly over small and relatively shallow structures, it has two serious disadvantages. The first is that the time taken to survey a given area is large: one measurement occupies about $1\frac{1}{2}$ to 2 hrs. The second is due to the fact that the gravity gradient is greatly influenced by local topographical features, and although a correction can be applied for this, errors in its determination lead to uncertainties in the gravity values deduced from the observations. Although the modern gravity meter attempts to replace the torsion balance for certain problems, it is of interest to compare their relative sensitivities. The torsion balance operates because there is a gravity change over its beam system, which is usually of the order of 40 cm. long, the gradient being calculated on the assumption that the change is uniform. A standard instrument will respond to a gradient of 10^{-9} sec.$^{-2}$ (1 Eötvös unit), which corresponds to a change of $4 \cdot 10^{-5}$ milligal over the beam. As will be seen later, gravity meters under laboratory conditions have been able

to detect variations of the order of 10^{-2} to 10^{-3} milligal, while a field instrument capable of measuring to the nearest 0·1 milligal would be regarded as satisfactory.

It is more useful to compare the accuracy of modern instruments with the accuracy of pendulums which measure the same entity. Accurate gravity observations date from the early nineteenth century, when Kater first described his reversible pendulum for absolute measurements. The inconvenience of these observations led to the introduction of invariable pendulums, with which gravity differences between stations are deduced from the difference in period of the same pendulum at the stations. This system, using pendulums of the Von Sterneck pattern, has proved most reliable, and the greatest number of observations have been made with it, the measurements being made with respect to some datum station at which the absolute value is known.

A big advance in the technique of field observations with pendulums has recently been made by Bullard (1933). In practice, it is usual to swing the field pendulum and a master pendulum simultaneously at a base station and again with one at the field station and the master pendulum at the base station. To observe the change in period it is obviously necessary for communication to be established between the two. With the advent of wireless, this allows widely separated stations to be compared, but it has the disadvantage that a powerful transmitter would be necessary to observe the difference in gravity between, say, Cambridge and a station in East Africa. Bullard made use of an already existing transmitter, viz. the Rugby Wireless Station, and recorded the same transmission at both Cambridge and the field station, and on the same charts records were taken of the swinging pendulums. As it was found possible to identify the same signal on both charts, a common time base was established between the widely separated stations. After applying the usual corrections, the accuracy was such that repeated values at the same station were in agreement to 1 or 2 parts in 10^6. One great advantage of the system is that, on account of improved methods of measuring the chart, the time of observation was reduced to 1 hr. This technique has been used for surveys in East Africa, Cyprus and elsewhere with considerable success (Bullard, 1937 and 1939).

Under laboratory conditions, however, a much higher order of accuracy has been obtained by Loomis (1931), who was able to measure the variation of gravity arising from the moon. For a rigid earth at latitude λ, this has an amplitude given by

$$g_m/g = 7{\cdot}6 \,.\, 10^{-8} \cos^2 \lambda,$$

where g_m is the gravitational effect of the moon and g that of the earth. The observations were carried out at Tuxedo Park in America ($\lambda = 41°$), where the fractional change is $4{\cdot}3 \times 10^{-8}$. To attain the necessary accuracy, the rates of three pendulum-operated Shortt clocks were compared with that of a quartz oscillator maintained under steady physical conditions during the experiment.

The oscillator had a frequency of 10^5 c./sec., and its output was used to control the frequency of a 1000-cycle generator. The current from the latter was fed

to a Rayleigh wheel which was maintained at a steady rate of 10 rotations per sec. Rigidly attached to the axis of the rotor was a bakelite arm carrying at its end a small tungsten plate B (figure 1). This swept round close to a fixed insulating ring on which were mounted, on the inner surface, a continuous metal ring A and 100 needle points C equally spaced around the periphery. Each needle point was connected electrically to a corresponding needle on a " comb " which was mounted with its teeth close to an earthed metal roller over which the recording paper was fed. At the end of each 30 swings made by the pendulums, a relay was operated, and this raised the continuous metal ring to a high potential, whereupon a discharge took place via the tungsten plate on the bakelite arm, from the ring to the nearest needle point, and a hole was punctured in the paper immediately

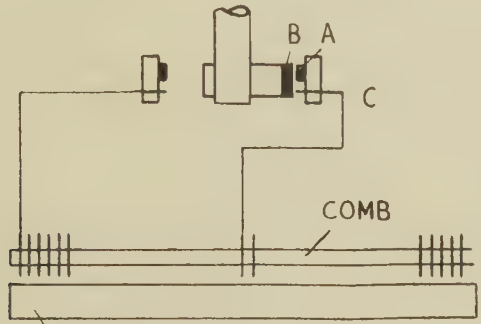


Figure 1.

below the corresponding tooth in the comb. It is obvious that, if 30 swings corresponded exactly to 300 rotations of the Rayleigh wheel, the record would consist of a line parallel to the edge of the chart ; an inclined line indicated that the pendulums were losing or gaining on the oscillator. Observations were carried out over 146 days, and an analysis of the results (Brown and Brouwer, 1931) was made for the whole period and for a 54-day period. It was found that the difference in the rates of the three clocks gave a random distribution with time, but the difference between each clock and the oscillator exhibited a periodic variation corresponding to half the lunar day. The values deduced from these observations agreed well with the theoretical value for a rigid earth.

§ 2. STATIC GRAVITY METERS

For geophysical prospecting, the pendulum suffers from one of the disadvantages of the torsion balance in that the time to complete an observation is long ; for this reason the static gravity meter is preferable. The conception of this type of instrument is due to Herschel (1833), who suggested the use of a spring, mounted in a metal frame and supporting a mass. The mass was to be varied until the total length was a constant, the metal frame being used as a gauge. Such an instrument is obviously crude and has many disadvantages for precise work,

but nevertheless many modern instruments consist of this simple system, and a few of them, e.g. the Hartley gravity meter, are used as null instruments.

At one time the need for a reliable gravity meter was urgent, to accelerate and supplement the numerous pendulum surveys which were then taking place, and in 1888 the British Association for the Advancement of Science offered a prize for a gravity meter complying with certain specified requirements, the most important being portability and an ability to measure gravity changes to the nearest 10 milligal. The award was never made, and a year or so later it was withdrawn pending the investigation of the elastic properties of the quartz fibres then newly discovered by Boys. It is interesting to note that the first gravity meter satisfying the conditions laid down by the British Association, viz. the Threlfall and Pollock instrument, did, in fact, use a quartz fibre as the elastic member (Threlfall and Pollock, 1900).

Judged by modern requirements, the standard laid down by the British Association is low, but even with the experience of modern instrument design and the greater number of materials available for their construction, it is by no means an easy matter to produce a gravity meter having an accuracy better than 1 milligal which at the same time is robust and sufficiently transportable to be employed in the field.

In static gravity meters, the force operating on a constant mass is opposed by some other force independent of gravity, such as elastic forces, magnetic or electromagnetic forces, etc. Of these, the first has proved most fruitful. The use of magnetic forces has not received much attention, for they usually vary rapidly with temperature and, more serious still, they would respond to the earth's magnetic field as the instrument position was changed, unless almost perfect magnetic shielding were employed. The last type of force immediately suggests an instrument taking the form of a modified Kelvin current-balance, but here there are considerable practical difficulties in the realization of a field instrument. The accuracy required leads to a beam system having the precision of a delicate balance in spite of the control exerted by the leads to the current coils mounted on the beam itself. In addition, as the applied force depends on the square of the current, field potentiometers capable of an accuracy of 1 in $2 \cdot 10^6$ would be necessary for the measurement of a milligal. Nevertheless, the force between two current-carrying coils may be found useful as a fine control in instruments using a null system of observation, or as a method of applying small known forces for calibration purposes. Here the accuracy of the current measurement need not be so great. In a null instrument the method would have two advantages: the current adjustment could be made on controls separate from the mechanical system, and a large range of gravity values could be covered without the necessity of modifying the main balancing system.

Gravity meters which depend on elastic forces can be divided roughly into two groups, those which depend on the direct application of Hooke's law and those which operate close to a position of instability to obtain greater sensitivity,

the latter including the so-called astatized instruments. As will be seen later, this division is not clearly defined, but it has advantages for the purpose of this paper.

Many of the factors which have to be considered in the design and construction of gravity meters are revealed by an analysis of a simple instrument consisting of a mass m suspended from a spring, leading to an extension x in a gravitational field g. Then $mg = Ex$, where E is the elastic constant of the spring. The sensitivity is given by $\Delta g/g = \Delta x/x$, where Δx is the additional extension for a gravity increase Δg. Thus, to measure a milligal, or less (since x can only be of the order of a few centimetres), considerable magnification is necessary. Again, the value of E changes with temperature, and there will be a change in the position of the mass due to this cause. If β is the temperature coefficient of the spring, it is easy to show that a temperature change of 1° C. will lead to an apparent change of gravity given by $\Delta g/g = -\beta$. As the temperature coefficients are of the order of 10^{-4} per 1° C., corresponding to 100 milligals, precise temperature control of the instrument, temperature compensation, or both, are essential.

It is well known that, if at constant temperature, an elastic body is taken through a loading and unloading cycle, a hysteresis loop is formed when deflection is plotted against the load. Accordingly, the deflection is not a unique function of the load, but may take up any value within the limits of the bounding curve. There is thus a limit to the accuracy of any spring gravity-meter imposed by the imperfections of the materials employed. For any given reading there is a range of gravity values within which the correct value must lie, and changes smaller than this range cannot be detected with certainty. This range naturally depends on the material, and for instruments capable of detecting less than 1 milligal there are only a few suitable, of which quartz, elinvar, certain steels and gases have been employed. Under a steady load, most springs exhibit a slow temporal change leading to a continuous drift of the readings. If the drift is not too great, it does not materially influence the accuracy of any survey, since, if readings at one or more stations can be repeated at suitable time intervals, a correction can be applied. With matured springs the observations usually lie on a straight line when plotted against time, and the excess over, or deficiency from, this line at the time of observation gives the change in gravity. One final effect due to elastic imperfections may arise from the clamping of the system for purposes of transport. If clamping introduces any serious strain, the system when released will move slowly to its final equilibrium position. If the instrument is read a suitable time after releasing, any error due to this cause will be negligible, but by correct design of the clamping mechanism the error can be eliminated.

In addition to the above, there are certain obvious precautions necessary to eliminate errors arising from changes in the buoyancy of the air, condensation and surface corrosion, etc. By selection of materials, any influence due to the earth's magnetic field can be eliminated as well as the effect of any electrostatic charges,

In one instrument the last has been avoided by incorporating some radioactive salt within the instrument. Again, a mass mounted on a spring, the system having a long period, forms a seismograph and, with the magnification usually necessary in gravity meters, the ground unrest is readily perceptible unless its effects are reduced by excessive damping.

Many instruments are very susceptible to errors in levelling, particularly those in which the moving parts are constrained to one degree of freedom relative to the instrument frame. Correct levelling ensures that the direction of gravity acts in the same direction as the only possible motion of the suspended system, while with incorrect levelling the appropriate component of gravity is operative. For an error of tilt θ, the effective value of gravity is changed by $\Delta g = -g.\theta^2/2$, and for an accuracy of 1 milligal the level must be reproduced to within $5'$ of arc. When possible, this effect is usually made useful for calibration purposes.

Although these factors have been discussed in relation to a simple spring instrument, similar considerations apply to most gravity meters, but naturally their relative importance depends to a great extent on the actual design. Thus, for example, in certain instruments, such as the Ising and Holweck-Lejay gravity meters, there is no permanent distortion of the elastic members, and accordingly there should be no drift correction of the type described. Nevertheless, a correction is necessary due to the ageing of the elastic member which results in a slow change of its constant.

§ 2.1. *Direct instruments with magnification*

There are a large number of instruments in this class, but only five will be considered here, viz. the Threlfall and Pollock gravity balance, the Hartley instrument, the Gulf gravity meter, the Boliden instrument due to Lindblad and

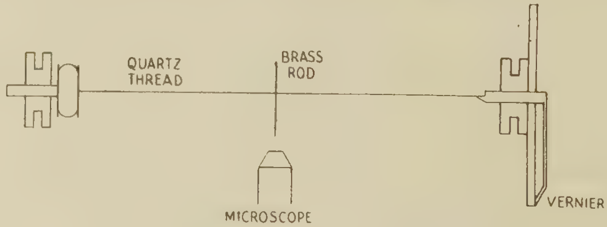


Figure 2. The Threlfall and Pollock instrument.

Malmqvist and one due to Haalck. Of these, the first is mainly of historical interest, but the others are modern instruments of greater sensitivity and are representative of present design.

(a) *The Threlfall and Pollock quartz-thread gravity balance.* This instrument (Threlfall and Pollock, 1900), which is shown diagrammatically in figure 2, consists essentially of a horizontal quartz thread, about 30 cm. long and 0·038 mm. in diameter. At the centre of the thread, and perpendicular to it, is attached

a gilded brass wire, of mass 0·018 gm., with its centre of gravity to one side of the thread. The ends of the thread are joined to rotating members whose axes are extensions of the quartz thread, the latter being kept taut by a spring inserted at one end. With no twist in the two sections of the quartz, the rod would hang vertically downwards, but rotating the ends in the same direction causes the rod to deflect from the vertical and, in the actual instrument, a rotation of 6π radians was found to bring the rod into its horizontal working position, where its end was observed by a fixed microscope. With the rod horizontal, the instrument is most sensitive to changes in g, the sensitivity being obtained by using a large initial deflection and the optical system. The direct method of observation was not employed, and the rod was restored to a standard position at each station by a rotation of one end of the thread which was attached to a sextant carrying a vernier. With the particular initial distortion, there exists a position of instability for small rotations θ of one end, when the rod is about 3° above the horizontal. Near to this the adjustment can be made with considerable accuracy. On theoretical grounds

$$\theta = Kg + C,$$

where K and C are instrument constants, and experience showed that the effect of temperature (t) could be accounted for by the modification

$$\theta = Kg(1+\beta t) + C.$$

The temperature was measured by a platinum thermometer supported on a thin-walled glass tube lying close to and parallel with the quartz thread. Calibration was performed using the known values of gravity at Melbourne and Sydney, and it was found that 2 minutes on the sextant corresponded to a change of 10 milligals, the limit of accuracy with this instrument.

(b) *The Hartley gravity meter* (Hartley, 1932). A simple stretched spring, which is adjusted to a constant length at each station, forms the essential part of the Hartley instrument. The greater part of the load is carried by the main spring (figure 3), only about 0·1 per cent being sustained by the lighter control spring. Both springs are hinged to a horizontal beam which carries the mass, and is attached at one end to the instrument frame by a metal filament about which the beam pivots. The upper end of the control spring is fixed to a micrometer screw which is used as the fine adjustment to restore the beam to the standard position. The main spring, of a tantalum-tungsten alloy, is stressed in the operating position to less than a fifth of that necessary to reach its elastic limit. As one end of the beam rises or falls under gravity variations, two mirrors mounted on light metal filaments are rotated in opposite directions, and the position of the system is observed by the separation of the two images of a lamp filament formed by reflection at the mirrors. The adjustment consists in bringing the two images into coincidence. The displacement magnification is about 60,000, but as part of the increase in the force of gravity is used to bend the various filaments, etc., the instrument has a somewhat smaller magnification.

Temperature compensation is attained by the usual gridiron system, but it was found essential to employ temperature regulation as well. Without this, as the temperature of the system is never uniform to the requisite degree of accuracy, distortion of the instrument and frame results. It is worthy of note that trouble was experienced from the initial strain of the frame and, as far as possible, it was

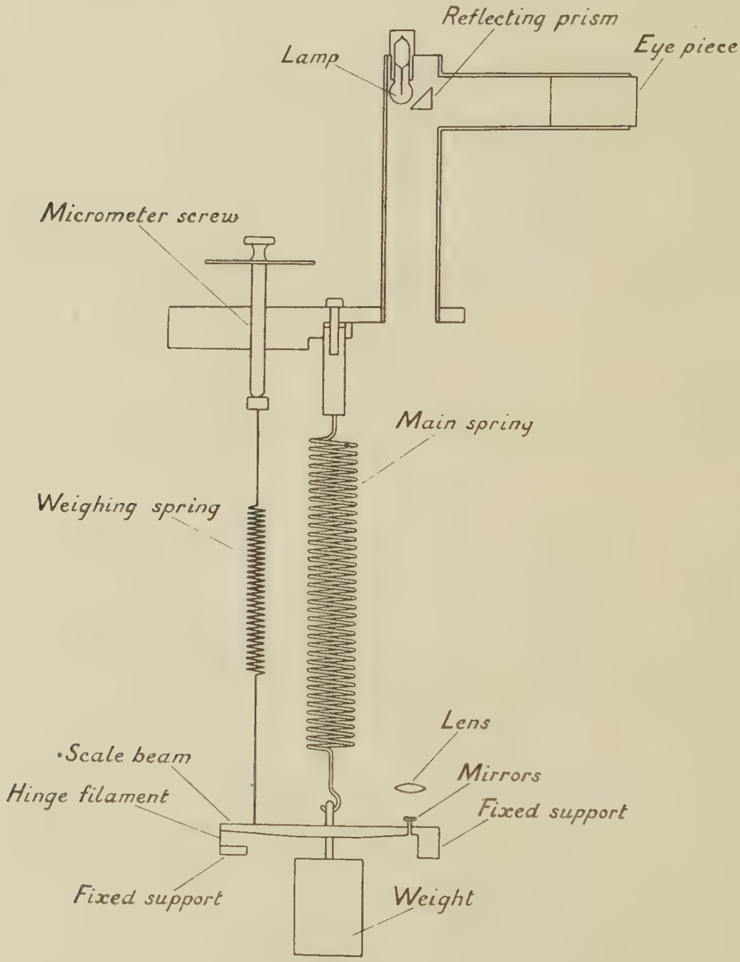


Figure 3. The Hartley instrument.

necessary to mount the optical system in a separate frame. The connection between this and the main frame was by annealed aluminium tubes fitting in accurately ground sockets and clamped by a method designed to put a minimum strain on the tubes. The most serious trouble arose from elastic hysteresis—not only of the main spring, but also of the control spring and filaments. To reduce this, the clamping system was designed to lock the beam in the standard position

without additional stress. Thus, it is claimed that the maximum stress, to which the system is exposed, corresponds to the gravity difference between stations.

The seismograph action is eliminated by excessive air damping. Finally, inaccurate levelling causes lateral distortion of the springs and filaments, in addition to errors arising from the effective component of g. Experiment showed, however, that with levelling errors less than 30″ of arc, gravity changes could be observed to the nearest milligal.

(c) *The Boliden gravity meter* (1938). This instrument was developed in Sweden by Lindblad and Malmqvist to be used in conjunction with those electrical methods of prospecting whose object is the location of good conducting zones. A number of geological formations, such as clays, graphitic schists, etc., behave as good conductors, and cannot be distinguished electrically from the denser ones such as galena etc. It was anticipated that a sufficiently sensitive gravity meter would allow a separation of the important and unimportant indications on account of their density difference. The mass (figure 4) is supported from the instrument

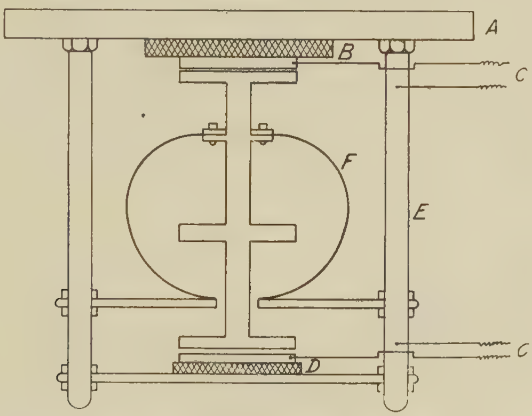


Figure 4. The Boliden instrument.

frame by two bent springs, and each end of the mass forms an electrical condenser with a metal plate insulated from, but rigidly attached to, the instrument frame. If a gravity change produces a displacement Δx of the mass, the corresponding change ΔC in the capacity C of the condenser is given by

$$\Delta C/C = -\Delta x/t,$$

where t is the separation of the plates. By making t sufficiently small, considerable magnification is possible. In the original paper, it is claimed that a movement of $3{\cdot}5 \times 10^{-9}$ cm. can be detected, while a change of 1 milligal causes a displacement of $5{\cdot}5 \times 10^{-7}$ cm. Thus, theoretically, gravity differences can be measured to the nearest 0·02 milligal. It is very doubtful if this accuracy is actually attained. The condenser forms part of an oscillating circuit whose frequency changes with the capacity, and the measurement is essentially one of comparing the frequency with that of an independent oscillator. Measurements

can be made by a null method by applying a known voltage between the plates and using the electrostatic attraction to restore the frequency to a standard value, a method which can also be used for calibration purposes. As this instrument only detects movements in one direction, i.e. perpendicular to the condenser plates, the method of calibration by tilting can also be adopted. By suitable selection of the materials used in its construction (these are not specified in the description), it is claimed that the effect of temperature can be made small, 1° c. change producing a movement corresponding to 2 milligals. In addition to temperature compensation, the temperature is regulated by thermostats to 0·02° c., and temperature effects can be ignored.

(d) *The Gulf gravity meter.* Details, apart from those in the relevant patent specifications, are not available for this instrument,* but it is included here as it employs a property of a helical spring, usually overlooked. It is shown in text-books (Poynting and Thomson, 1924) that a helical spring, as well as extending under

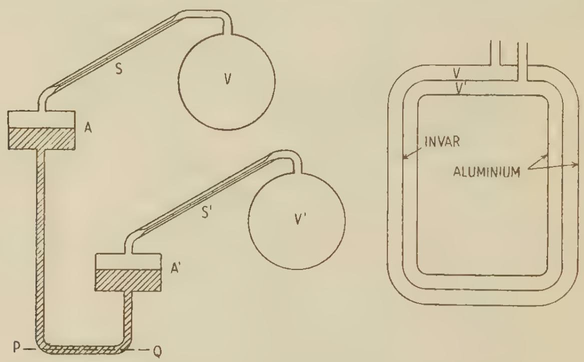


Figure 5. The Haalck gravity meter.

a load, suffers a twist of one end relative to the other. For a spring of circular wire the effect is small, but by suitable selection of the cross-section it can be accentuated. With a rectangular section and the long dimension parallel to the axis, the spring uncoils on loading; with the short dimension parallel to the axis the spring coils up. The Gulf instrument employs a spring of the former type, the rotation of the lower end being observed by a modification of the reflected light beam system. Using two partially silvered mirrors, multiple reflections are used, the displacement of the images being in arithmetic progression. A high-order reflection is observed, and when this moves off the scale, the next lower order is used. By this means a wide range of gravity values can be examined without adjustment of the suspended system.

(e) *The Haalck gravity meter* (Haalck, 1932). Many instruments have been proposed and constructed on the barometer principle, and the Haalck gravity meter is one of the latest designed on these lines. If a mass of gas is enclosed under the pressure of a column of mercury, when the system is moved to a point

* A description has now been published in *Geophysics*, 1941, 6, 13.

of increased gravity, the mercury column exerts a greater pressure and the gas is compressed, and attains an equilibrium under a shorter column. Under normal conditions, fluctuations in atmospheric pressure and temperature would swamp the gravity effects, and Haalck has made efforts to eliminate or reduce these extraneous effects, at the same time obtaining considerable magnification by the use of a double liquid manometer. The atmospheric pressure is immediately eliminated by measuring the differential pressure between two gas volumes V and V′ (figure 5). The main manometer fluid is mercury but, as in other instruments, the fractional change in pressure is only of the same order as the fractional change in gravity. The mercury surfaces are given large cross-sections A and A′, and toluene is used above the mercury with its surfaces in the capillary tubes with small sections S and S′, the capillary tubes being inclined at a very small angle to the horizontal. The magnification depends on the incompressibility of the toluene. It can be shown that

$$\Delta g/g = C\Delta l - C'\Delta l',$$

where C and C' are instrument constants and Δl and $\Delta l'$ are the displacements of the liquid in the capillary tubes. Of necessity these are of opposite signs, and Δg depends on the sum of the displacements. One advantage of the system is the ease with which the constants can be determined. A plunger allows a small quantity of mercury to be added to the manometer, compressing both gas volumes, and the displacements are related by

$$\Delta l/\Delta l' = C'/C.$$

In addition, the plunger allows the values to be adjusted to the necessary gravity range. Finally, by tilting the system about the axis PQ, all vertical distances are modified by the factor $\cos\theta$, where θ is the tilt. This corresponds to a reduction in gravity of $-g\,\theta^2/2$, allowing the individual values of C and C' to be determined.

The most important correction is for temperature, and as 1° c. corresponds to about 4000 milligals, temperature control must be supplemented by compensation. The temperature effect is complex, but the most important part depends on a term $\left(\alpha - 3\,\dfrac{\epsilon p - \epsilon' p'}{p - p'}\right)$, where α, ϵ and ϵ' are the temperature coefficients of the gas and the containing vessels, and p and p' are the pressures. One volume of gas is made to enclose the other, the common wall of the two vessels being of invar. The inner and outer walls are of materials having a large temperature coefficient, and by suitable selection of the dimensions the inner volume can be made with a large negative coefficient and the outer volume with a large positive coefficient. By this means the term can be made small, but perfect compensation cannot be attained, and melting ice is used to maintain a constant and uniform temperature.

Haalck reports that, with a set of four such instruments mounted in a van, an accuracy of 1 milligal can be attained by using a repeated set of observations. As the manipulation of the instrument is so simple, he suggests the possibility

of making four or five repeated observations, leading to an accuracy of 0·25 to 0·5 milligal, but it should be noted that sporadic deviations of 3 milligals have been noted for a single observation.

Another instrument in which gas is used as the elastic body is that due to Nørgaard (1933, 1935, 1936). It takes the form of an inverted Cartesian diver in which a gravity change alters the pressure and volume of the mass of gas enclosed in the float. This accordingly moves to a new position of equilibrium where the total volume of liquid displaced by the float and gas is the same as before. Recently, Benfield (1938) has attempted to construct an instrument on this principle using only mercury as the manometer fluid instead of a combination of two fluids as in the case of Nørgaard's instrument. With the sensitivity aimed at, i.e. 1/20 mm. displacement for 1 milligal, the instrument was unsuccessful, mainly on account of the irregularities in the surface tension of the mercury and the contact angle between it and the float.

§ 2.2. *Instruments operating near a position of instability*

Only three of the many such instruments will be described, one, due to Thyssen (1939), being a simple extension of the spring-balance, and the other two due to Ising and Urelius and Holweck and Lejay. The last do not suffer from the great disadvantage of all the other instruments described here, for there is no permanent distortion of the elastic member. Both employ the same physical ideas, but the method of observation and the instrument construction differ considerably.

(a) *The Thyssen gravity meter* (Thyssen, 1939). A quartz beam (figure 6), resting on a central knife-edge, has a platinum mass of 20 gm. at one end, the whole being held in equilibrium by the spring, some 50 cm. long, attached to the other end. The beam is designed so that its centre of gravity is just below the axis of rotation. A small mass of 1 gm., moving on a screwed thread mounted on the centre of the beam and above it, allows the centre of gravity to be raised until an unstable position is attained. In practice the upper mass is adjusted until the system has a period of 6 to 10 sec., and the position of the beam can be controlled at the standard base station by a micrometer at the lower end of the spring. The deflections are observed by a prism and microscope (magnification $\times$ 60) focused on a vertical scale fixed to the main platinum mass.

The spring itself is of a nickel-iron alloy, which by suitable annealing may be obtained with a wide range of elastic temperature coefficients from about $-25 \,.\, 10^{-6}$ to $-10 \,.\, 10^{-6}$. The thermal expansions of a number of materials lie within this range (but are of opposite sign), and temperature compensation is effected by selecting for the tube surrounding the spring a substance whose expansion coefficient has the same magnitude as the temperature coefficient of the spring. Perfect compensation is not possible, but the error can be reduced to a value that allows a correction to be applied, the essential condition being that the spring and tube shall be at the same temperature. This part of the instrument is heavily

lagged for this purpose. Temperature regulation is not employed, in complete opposition to the experience of Hartley, who found that, without it, non-uniform temperatures led to distortions which vitiated the readings. In addition to heavy lagging round the spring tube, the whole is covered with felt as an additional precaution.

It is obvious that the beam system will be sensitive to tilt about an axis parallel to the knife-edge. Errors from this cause are eliminated by using two identical beams side by side, but with their main masses at opposite ends. Small errors in levelling, which is carried out by a 5-sec. bubble level, have equal and opposite effects on the beams, and the reading for a true adjustment can be calculated.

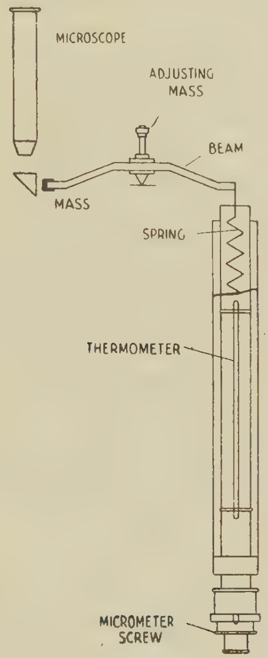


Figure 6. The Thyssen instrument.

The use of a quartz beam renders the instrument susceptible to electrostatic charges, and these are avoided by incorporating a little radioactive salt in the case. It should be noted in this connection that errors are not only due to the interaction between charges, but also to the attracted dust particles, which alter the mass distribution.

(*b*) *The Ising and Urelius gravity meter* (Ising and Urelius, 1930). In construction, the instrument is reminiscent of the Threlfall and Pollock gravity meter, and consists of a horizontal thread of quartz KL (figure 7 (*a*)) fused at each end to a quartz stand Q, and at its centre is fixed a quartz rod P with its centre of gravity displaced from the thread. The normal equilibrium position of the rod is vertical and above the thread, and in this position there is no twist in the fibre. If the

effective elastic constant τ of the thread is greater than mgh (m being the mass of the rod and h the distance of the centre of gravity from the thread), the upright position is one of stable equilibrium, for any deflection θ results in a deflecting couple $mhg \sin \theta$ and a restoring couple $\tau\theta$, the latter always being the greater. These two are shown in figure 7 (*b*) by OCD and OF respectively. The instrument is constructed so that the elastic constant is only just greater than mgh, so that the

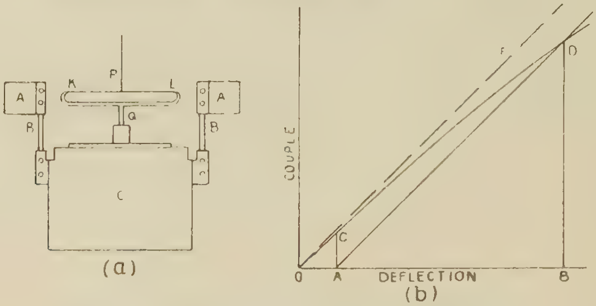


Figure 7. The Ising and Urelius gravity meter.

graphs are nearly parallel. In operation, the whole instrument is given a small tilt θ ($=$OA), when a gravitational couple AC operates without any twist in the thread. The rod will accordingly rotate through an additional angle ϕ until equilibrium is attained represented by the conditions D. Since the two lines are nearly parallel ϕ is very much larger than θ. It can be shown that

$$\phi/\theta = mgh/(\tau - mgh) = N,^*$$

and it is the variations in N which allow the gravity changes to be observed, for N obviously changes rapidly with g. If N and $(N+\Delta N)$ correspond to the gravity values g and $(g+\Delta g)$ then, approximately,

$$\Delta g/g = \Delta N/N^2.$$

Thus, for example, if N is about 1000, measurements of N with an accuracy of 1 in 1000 yield gravity differences to the nearest milligal.

An instrument of this nature is obviously sensitive to faulty levelling, and to reduce errors arising from this the quartz stand is mounted on a heavy lead block C supported from the instrument frame A by vertical laminar springs B. After levelling the frame, any inaccuracies are automatically eliminated by the yielding of the springs. In the early instruments, the position of the rod was observed by a microscope, the objective being on the block and the eyepiece on the frame. The small tilt θ was obtained in one of two ways, either by loading the mass with a rider placed at one side, or by means of a light spring attached to the lead block with its tension controlled by a fine-pitched thread. In the former case the deflections for a constant tilt were obtained, and in the latter the tilt necessary to produce a constant deflection was observed. The temperature was maintained

* In practice, it is necessary to take into account the fact that the gravitational couple is proportional to sin $(\theta+\phi)$ and not to $(\theta+\phi)$ itself.

constant by an ice bath, but in a later model thermostatic control has been used. Automatic recording has been developed for the load method of taking readings, and, using this system, experiments in England and Holland gave a mean error of 0·46 milligal, while the visual screw method of observation gave 0·64 milligal.

(c) *The Holweck-Lejay inverted pendulum* (Holweck and Lejay, 1934). Although the same physical principles are used in this instrument as in the Ising-Urelius gravity meter, their practical forms differ considerably. The essential parts of the former consist of a rod of quartz D (figure 8) mounted on a vertical

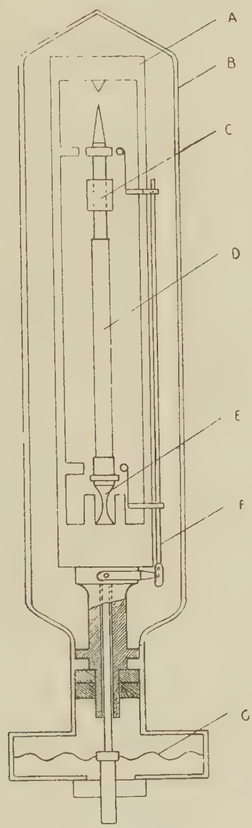


Figure 8. The Holweck-Lejay pendulum.

spring E whose lower end is clamped in the frame F. The elastic constant τ of the spring is again adjusted to a value slightly greater than mgh. It can be shown that, under ideal conditions, the period T of the pendulum is

$$T=2\pi\sqrt{I/(\tau-mgh)},$$

where I is the moment of inertia of the rod, and hence

$$\Delta T/T=mh\Delta g/(\tau-mgh).$$

Comparing this with the corresponding equation for an invariable pendulum,

$$2\,\Delta T/T=\Delta g/g,$$

it will be seen that the sensitivity has been increased by a factor $mgh/(\tau-mgh)$, which is obviously large. This magnification can be attained experimentally by observing the period (T_1) with the pendulum inverted and with the pendulum hanging downwards (T_2). In the latter position, gravity and the spring assist one another, leading to a short period. Then

$$T_1=2\pi\sqrt{I/(\tau-mgh)}\,;\quad T_2=2\pi\sqrt{I/(\tau+mgh)}$$

and

$$2mgh/(\tau-mgh)=(T_1^2-T_2^2)/T_2^2.$$

In one particular case $T_1=6{\cdot}3477$ sec. and $T_2=0{\cdot}339$ sec., giving

$$\Delta T/T=87{\cdot}4\,\Delta g/g.$$

Thus for an accuracy of a milligal, T must be observed to one part in 10^4, and as observations are usually made on 100 oscillations, the timing must be carried out to 0·06 sec.

Because the pendulum is made sensitive to changes in gravity, it is also sensitive to other extraneous effects, notably to tilt and to amplitude. For the above pendulum with an amplitude of 1°, the correction is 33.10^{-4} sec. Normally, the pendulum is started with the same amplitude, so this correction is constant and, due to damping, somewhat smaller than this. The correction for tilt depends roughly on the square of the amplification and is extremely important. The adjustment of the level is carried out by making use of the fact that the pendulum gives a magnified indication of the tilt as in the Ising instrument.

By suitable selection of materials, errors arising from temperature changes can be reduced to small values, and during observation the temperature is observed and a correction, based on experimental data, is made: a correction is also made for the temporal change in the elastic constant of the spring, and this is found to vary according to the relation

$$C=C_\infty(1-10^{-\alpha(t-t_0)}),$$

where the notation is obvious. In the early life of the pendulum this effect is important, but ultimately becomes small.

The pendulum consists of a fused-quartz cylinder D, about 6 cm. long by 4 mm. in diameter, mounted on an elinvar spring E. The spring, which is worked in one piece, has massive ends where it is attached to the quartz and the frame, but is thinned in the middle. The period, and consequently the sensitivity, can be varied by means of an adjustable quartz ring C sliding on the rod. Electrostatic action is eliminated by coating the quartz with platinum and placing the whole in a Faraday cage. The pendulum is observed by a microscope focused on a fine quartz thread at the upper end of the rod. A thermometer has its bulb in a massive metal frame A which surrounds the pendulum and assists in maintaining a uniform temperature, while atmospheric conditions are eliminated by sealing the whole in a glass vessel B. It has been found necessary for the observer

to be insulated mechanically from the instrument, as it is susceptible to ground motions.

Instruments of this class suffer from the disadvantage that their sensitivity varies as the value of gravity changes, and they can only be used over a small range for which the instrument is adjusted. With increasing gravity the system ultimately becomes unstable and for decreasing gravity the sensitivity falls. In addition, as they are susceptible to errors in levelling and, in the case of the Holweck-Lejay instrument, to ground vibration, they do not appear suitable for observations at sea.

Reviewing the modern instruments, of which a representative selection has been described, it is apparent that under field conditions they are capable of measuring small changes in gravity with an accuracy somewhat better than the invariable pendulum under the same conditions, and in general the time of a single observation compares favourably with the time of pendulum measurements. In the laboratory, the static instruments have also been used to observe the gravitational effect of the moon. Amongst others, Tomaschek and Schaffernicht (1933) have successfully carried out such experiments. Using a simple loaded spring and an interferometer method of observing displacements, the moon was found to give a shift of about two fringes, while by using a bifilar gravity meter designed by Berroth, near a position of instability, a deflection of 2 mm. was obtained for $10^{-8} g$.

DISCUSSION

Prof. A. O. Rankine. Dr. Bruckshaw has repeated the widely believed statement that fused quartz does not display hysteresis effects. This is not true. In the Ising gravity meter which we have heard described, the marked hysteresis of the quartz has to be allowed for by making a series of observations at regular intervals. Liability to persistent electrification is another difficulty with quartz suspended systems. This also is recognized in the Ising instrument, arrangements being made for ionizing the air in the suspension chamber.

I agree with Dr. Bruckshaw in his quotation from Hartley that experience shows it to be impracticable to produce a working gravity meter measuring to one milligal merely by compensating temperature effects—a thermostatically controlled enclosure is essential. It seems to me desirable formally to rebut claims to the contrary, emanating chiefly from Germany. The work and claims of Baron von Thyssen are a case in point. A good deal of useful foreign exchange has found its way into the Reich through the unscrupulous exploitation outside Germany, with the backing of the Nazi Government, of the notoriously expensive but unreliable Thyssen gravity meter.

I am sorry Dr. Bruckshaw could not spare time to say more about the Gulf gravity meter. With a fairly wide experience of gravity meters I am firmly convinced that this instrument, depending on the rotation of the end of a suitable spring under varying load, and not upon its change of length, is much superior to any of its contemporaries.

Mr. Lancaster-Jones. The rapid development of gravimeters in recent years affords substantial support for a contention that I used to make in regard to the magnitudes measured by the Eötvös torsion balance, namely, that the gradient values were of prime significance and the curvature values of far less importance. The reason was that interpretation of curvature values is very difficult, whereas the gradients, when integrated into isogams, i.e. gravity-difference curves, are readily interpreted.

In stating the reason for the impetus to produce gravimeters, Dr. Bruckshaw has emphasized two factors: the saving in time, and the avoidance of corrections due to irregularities of density in the surroundings. Whilst I would fully endorse the influence of the former factor, I would very much doubt the value of the latter. In fact, it is a question of the nature of the gravity effects to be determined. In many problems, the values sought, due to the presence of some deposit of economic significance, are far smaller in amount than any gravimeter can detect. In some work carried out by Dr. Shaw and myself in Cumberland, concerned with the detection of comparatively large deposits of iron ore, the total gravity anomaly in a region covering a square mile or so was very little more than one milligal, and this effect was almost wholly due to undesired subterranean density anomalies; the effect of the ore-body itself was very much less.

Moreover, whilst gravimeters are indeed less sensitive to very local irregularities of density, they are more sensitive than torsion balances to regional irregularities in the topography, etc., because they operate over larger areas. In the same way pendulum measurements need to be corrected for large-scale topographical and other terrestrial effects. Indeed, as in most problems, the instrument must be adapted to the relative magnitude of the task and the field of operation.

Dr. H. Shaw. Dr. Bruckshaw at the outset of his lecture divided gravimeters into two groups, according to whether they were designed to measure gravity at sea, or for use in the location of minerals of economic importance. I am afraid that such a grouping overlooks certain early instruments which were devised for use on land, and were not capable of operation at sea. In order to include these instruments I would suggest that it might be preferable to re-name his first group—" Gravimeters designed for purely geodetic purposes ".

I was pleased to hear Dr. Bruckshaw describe the apparatus of Threlfall and Pollock as the earliest gravimeter to possess modern sensitivity and accuracy. It may be of interest to record that this early apparatus is now at the Science Museum, S. Kensington. When Sir Richard Threlfall brought it there, in 1932, he stated that for some years he had intended to continue his investigations. He discovered however that he had unfortunately lost his earlier manipulative skill, and had not been able to persuade anyone to continue the work in which he was so keenly interested.

Following upon the remarks of the previous speakers, it may be as well to

emphasize a distinction between the pendulum and the gravimeter, which I do not think has yet been mentioned. The gravimeter is an ideal instrument for making determinations of gravity at stations separated by distances up to several kilometres, where observations can be made at intervals up to, perhaps, one or two hours. The comparison of gravity at well-separated stations, however, such as London and New York, cannot be carried out with a gravimeter owing to certain errors inherent in this type of instrument. For the comparison of gravity at such well-separated stations some form of pendulum must be employed which at the same time enables an absolute determination of gravity to be made.

Dr. J. H. JONES. Dr. Bruckshaw will be interested to know that an account of the Gulf gravimeter has now been published in the latest number of *Geophysics*. There is no doubt that this is a very fine instrument, and its success is due chiefly to the neat method of magnifying the change in length of the spring. There is only one point of contact to the frame, and in consequence the instrument is insensitive to small levelling errors.

It is my experience that the two most important problems in gravimeter design are (1) the elimination of the effects of clamping and (2) the elimination of levelling sensitivity.

It is necessary to clamp the instrument as close as possible to the equilibrium position to eliminate strain on the spring. The elimination of levelling sensitivity is difficult if there are more points of contact to the frame than one, but in general this can be achieved by introducing constraints so that the moving system has only one degree of freedom.

REFERENCES

BENFIELD, 1938. *Mon. Not. R. Astr. Soc., Geophys. Suppl.* **4**, 351.
BROWN and BROUWER, 1931. *Mon. Not. R. Astr. Soc.* **91**, 575.
BULLARD, 1933. *Proc. Roy. Soc.* A, **141**, 233.
BULLARD, 1937. *Mon. Not. R. Astr. Soc., Geophys. Suppl.* **4**, 114 ; 1939. *Ibid.* **4**, 473.
HAALCK, 1932. *Z. Geophys.* **8**, 17, 197.
HARTLEY, 1932. *Physics*, **2**, 123.
HERSCHEL, 1833. *Outlines of Astronomy.*
HOLWECK and LEJAY, 1934. *J. Observations*, **17**, 109.
ISING and URELIUS, 1930. *Bull. Géodésique*, Annexe 7, 556.
LOOMIS, 1931. *Mon. Not. R. Astr. Soc.* **91**, 569.
NØRGAARD, 1933. *Balt. Geodet. Komm.* 6 Tag. 211; 1935. *Ibid.* 7 Tag. 279; 1936. *Ibid.* 8 Tag. 127.
POYNTING and THOMSON, 1924. *Properties of Matter*, 10th edition, p. 103.
SUNDBERG, 1938. *Bull. Instr. Min. Met., Lond.*, no. 402.
THOMASCHEK and SCHAFFERNICHT, 1932. *Z. Geophys.* **8**, 331.
THOMASCHEK and SCHAFFERNICHT, 1933. *Z. Geophys.* **9**, 125.
THRELFALL and POLLOCK, 1900. *Philos. Trans.* **193** A, 215.
THYSSEN, 1939. *Z. Geophys.* **15**, 121.

THE HYSTERESIS CYCLE AND ITS INTERPRETATION

By Professor L. F. BATES

Lecture delivered 25 *April* 1941

I BEGIN by reminding you of the various ways in which we may obtain a hysteresis cycle for a specimen of ferromagnetic material. The slides I now show illustrate the ballistic method, the magnetometer method and a method which does not involve the use of any instrument except a milliammeter. The latter is very suitable for dealing with rod-shaped specimens of highly permeable materials. It depends for its action on the Schuster and Smith method of measuring H and is fully described elsewhere (Bates, 1931).

I now show you a typical (B, H) curve for annealed iron-wire obtained by the late Sir Alfred Ewing in 1873. Many of the Fellows present will remember the last occasion on which Sir Alfred addressed the Society, and how he told us that he thought "no small beer" of himself when he first obtained such curves. In many respects Ewing's methods of describing the hysteresis cycle have not been altered, but with the increasing demand for materials of very high coercivity it is necessary to distinguish between the coercivity obtained from the (B, H) curve and the coercivity obtained from an (I, H) curve; this is well illustrated by a diagram in which a (B, H) curve and a $(4\pi I, H)$ curve are plotted together. The curves intersect at $H=0$, giving us the residual induction; but they do not pass through the same point on the H axis, and the coercivity obtained from the (B, H) curve can be considerably less than the coercivity given by the (I, H) curve. Users of permanent magnets, of course, specify the coercivity in terms of the (B, H) curve which may be written ${}_BH_c$. Ewing explained his results in terms of his well-known simple model of a large number of elementary magnets more or less freely pivoted, a model which must be replaced for modern purposes by a different picture of the whole process of magnetization, to which Becker and his collaborators have contributed much (Becker and Döring, 1939).

It is perhaps profitable to consider for a few moments a few of the hysteresis cycles which may be obtained in practice. First we have the enormous changes which can be produced in the magnetic properties of iron which is hydrogenized by heating the material in moist hydrogen at a very high temperature. There is an enormous increase in the initial permeability. Next we have the group of nickel-iron alloys whose magnetic properties depend so markedly upon their heat treatment, and we distinguish between air-quenched, annealed and baked

specimens. The air-quenched materials provide us with materials of very high initial permeability, of which *Permalloy* is a well-known example. Then we have that remarkable group of materials known as the *Perminvars*, alloys of iron, nickel and cobalt, which give the most remarkable set of hysteresis curves which man has yet seen. These alloys when "baked" exhibit the most important property of constant initial permeability over fields as great as 2 or 3 oersteds. It is necessary, however, to avoid exceeding these fields or a complicated hysteresis cycle will be obtained and the material will lose its important property.

Finally, we come to the magnetic properties of single crystals of ferromagnetic elements. In the case of iron the crystal is of the body-centred cubic type, i. e. the atoms are arranged one at each corner of the crystal cube and one in the centre of the cube. The magnetic properties of such crystals depend upon the direction in which the magnetizing field is applied, for in general I and H do not coincide in direction. We may distinguish between the three important directions of magnetization [100], [110] and [111]: it is found that magnetization takes place most readily when the field is applied in the [100] direction, which is called the *direction of easy magnetization*. The [111] direction is the direction of most difficult magnetization. Perhaps it is as well to point out that most of these curves have been obtained with very small specimens in which the demagnetization factor is of considerable importance, and there seems to be some diversity of opinion about the initial course of the curve and the positions of the marked kinks. In my opinion, such crystals do not have an infinitely great initial permeability, and even a perfect iron crystal would show some hysteresis phenomena. In the case of the single crystals of nickel, the [111] direction is the direction of easy magnetization while the [100] is that of difficult magnetization. The cubic crystal of nickel belongs to the face-centred type, i.e. with one atom at each cube corner and one atom at the centre of each face. It is clear that the magnetic properties of these crystals are of great theoretical significance, and should be borne in mind in the remaining portion of my talk.

We now interpret magnetization processes in terms of the Weiss *Domain theory* of magnetization. Weiss postulated the existence of a huge internal field in a ferromagnetic substance, in order to explain the magnitude of ferromagnetic phenomena and the existence of permanent magnetization. In recent years Heisenberg has explained the existence of this field on the basis of exchange forces between electrons. In the case of a hydrogen molecule, the atoms are bound together by a homopolar binding between the electron spins which are oppositely directed, so that the molecule is magnetically inert. In the case of ferromagnetic materials we have similar forces between electron spins which are in this case directed parallel to one another.

We must therefore picture a ferromagnetic as made up of a very large number of Weiss domains, whose size is, on the average, about 10^{-9} c.c. We consider that to a first approximation the magnetization, I_s or $I_{0,T}$, of a domain remains constant at a constant temperature and that its direction of magnetization, or the position

of the magnetic vector of the domain, alone alters when the domain is exposed to an applied field. In the accompanying diagram, figure 1, a number of such domains are shown. The boundaries between them are shown with a definiteness which, of course, we cannot imagine them to possess in practice. We will imagine for the present that each magnetic vector is parallel to a crystal axis, and we will now consider what might be expected to happen when we apply a field H in the direction shown. Firstly, the vector of domain 4 may turn through 180 degrees, so that while it still remains parallel to a crystal edge, it is in much better alignment with H. In like manner, the vectors of domains 2 and 3 might turn through 90 degrees. Secondly, the vector of domain 1 might grow at the expense of domain 2 as shown by the dotted line. Finally, after all the domain vectors have set themselves as nearly parallel to H as possible, although still remaining parallel to cube edges, they might thereafter rotate

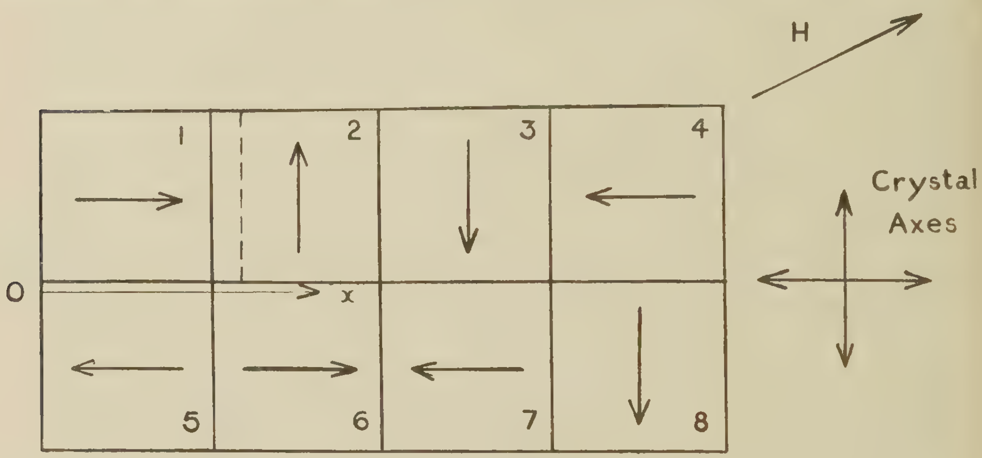


Figure 1. Illustrating domain structure of a ferromagnetic crystal.

bodily through angles less than 90 degrees, until they are exactly parallel to the field.

It is frequently assumed that all the spins within a given domain can simultaneously swing through large angles from one position of minimum domain-energy to another. The energy needed for such a comparatively large-scale process is too great to be drawn from the crystal lattice and the assumption is undoubtedly incorrect, so that the instantaneous 180-degree rotation of all the spins in domain 4 and a similar 90-degree rotation of all the spins in one of the domains 2, 3 and 8 is impossible. Consequently, the domain changes must proceed by means of a series of rapid rotations of individual spins. Moreover, while it is easy to draw domains as if they were of regular shapes situate within a two-dimensional frame, it is not at all easy without long experience to visualise their actual arrangement in three dimensions within a crystal. In any case, in strained materials the vectors will not make angles of exactly 90 and 180 degrees

with one another. Again, while we have thus far confined attention to single crystals, their properties are not unique, and all ferromagnetic materials must be regarded as collections of such crystals giving rise to hysteresis cycles which must be explicable in terms of the properties of the single crystals of which they are constituted.

Let us pause for a moment to consider the evidence for the existence of domains within a ferromagnetic body. We have first of all the Barkhausen effect, which I now propose to show you. On the table before me I have a coil of wire which is connected to a valve amplifier and loud-speaker. Within the coil is placed a piece of fine iron wire. On bringing a powerful permanent magnet up to this wire, the loud-speaker emits a series of discontinuous noises or clicks. If the magnet is brought up fairly quickly, then the noises approximate to those made by a body of recruits ordering arms for the first time. This phenomenon shows us that magnetization processes, particularly those on the steep portion of a Barkhausen curve, are essentially discontinuous, and each tiny portion of a smooth magnetization curve would, if suitably magnified, appear as a series of irregular steps. The phenomenon has been investigated in detail by many workers. Sixtus and Tonks (1931) have measured the velocity of propagation of large Barkhausen discontinuities along a stretched wire, for by suitable thermal and other treatment it is possible to produce very large nuclei which show big Barkhausen effects. Sixtus and Tonks have shown that it is possible to cause Barkhausen discontinuity to stop at a given point in a stretched wire and also to arrest its progress by the application of a suitable field in the reverse direction.

It is now known that there are three important factors which decide the direction of magnetization within a particular domain. Experiments with single crystals show conclusively the existence of directions of easy magnetization, e.g. in iron crystals parallel to the [100] directions, and in nickel crystals to the [111] directions. This is because different quantities of energy are associated with different positions of the vector with respect to the crystal axis, and we express this fact (Bates, 1939) by means of the anisotropy coefficients in the equation

$$E_s = K_0 + K_1(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) + K_2(\alpha_1^2\alpha_2^2\alpha_3^2).$$

Hence we may consider that the *crystal energy* per unit volume, E_c, is one of the factors deciding the position of the domain vector.

A second factor must be the effect of strain produced by the magneto-strictive change of dimensions which accompanies magnetization. It is impossible to prepare a ferromagnetic substance which is perfectly strain-free, because as it cools through the Curie point, the substance changes from a paramagnetic to a ferromagnetic state, resulting in stresses which cannot be entirely removed, however careful and prolonged subsequent annealing may be. We know from direct experiment how important such stresses can be; for example, a nickel wire under severe longitudinal stress can only with difficulty be magnetized parallel to its axis, because the domain vectors under these circumstances tend

to set perpendicular to the axis of the wire. In the case of hard-drawn nickel unstrained by external forces, they set preponderantly along the axis. In fact, it can easily be proved that if we apply a tension of F dynes per sq. cm. to a domain in nickel, for which E_c is small, we have to supply *strain energy* per unit volume, E_F, equal to $\frac{3}{2}\lambda_s F \sin^2\phi$, in order to turn the vector through an angle ϕ from the direction of the applied tension, λ_s being the saturation longitudinal magnetostriction coefficient.

Finally, in addition to E_c and E_F, we have the *field energy* per unit volume E_H, which is the energy of position or potential energy of the domain magnetization *with respect to the applied field*; it is not the energy located in the material itself, as we shall see later. This is clearly equal to $-HI_{0,T}\cos\theta$, where θ is the angle between the domain vector and H. Consequently, to sum up, the domain vector must set in such a direction that $E_c+E_F+E_H$ is a minimum. It follows immediately that only when the material is approximately strain-free does E_c decide the direction of the vector, and that only when E_H is greatly in excess of E_c and E_F can its direction be fixed uniquely. Now, there must always exist a minimum internal stress equal to $F_i=\lambda_s Y$, where Y is Young's modulus for the material, because the latter gains its ferromagnetism and is correspondingly distorted on cooling through the Curie point. It is therefore appropriate to divide ferromagnetic materials into two groups according as $E_c \gg E_F$ or not.

We may now give a somewhat more definite picture of what takes place when we magnetize a ferromagnetic substance like iron, for which E_c is always $\gg E_F$, where, in fact, tensions considerably above the elastic limit would be needed to make E_F approach E_c in magnitude. Let us consider two adjacent domains, such as 1 and 2 in figure 1, whose vectors are at 90 degrees with respect to one another and with what is termed a 90-*degree boundary* between them. Let us plot the difference in the energy minima per unit volume, denoted by E_1-E_2, as we proceed through the two domains in the x direction; i.e. we imagine unit volume around any given point to have its vector pointing first as in 1 and then as in 2. It will be represented by a curve of the form ABCDEF, figure 2, when no field is applied, so that the boundary is located by the point of intersection A on the x axis. On establishing an effective field H_1 parallel to x, the energy difference is decreased by $H_1I_{0,T}$, which may be regarded as equivalent to raising the x axis by this amount, so that it takes up the position B_1B and intersects the curve at B, which now represents the new boundary position. Proceeding in this way with larger values of the applied field, successive points of intersection representing a 90-*degree boundary displacement* are obtained. This displacement is magnetically reversible until we reach the point C, but once the field is increased beyond the value for C, a region is found where the x direction of the vector is particularly favoured, and the boundary moves forward rapidly to positions given by such points of intersection as E. The change between C and E is magnetically irreversible and forms a Barkhausen discontinuity. On decreasing the field the minimum D is reached before a similar Barkhausen change takes place in the reverse direction,

When two adjacent domains whose vectors are anti-parallel are under consideration, it is necessary to imagine a transition layer between them analogous to the layers considered in surface-tension problems. The experiments of Sixtus and Tonks (1931) show that a critical field must be exceeded if a more favoured domain is to grow at the expense of another. This means that such changes are irreversible. They undoubtedly account for the majority of the Barkhausen discontinuities, particularly in the region below the knee of the hysteresis curve.

Now, the fields, necessary to cause the two types of boundary displacement to continue until all the domain vectors in the iron crystal are set parallel to the cube edge most nearly coincident with the direction of the applied field, are not large. If it is further desired to turn the direction of magnetization from parallel to the cube edge through an angle of less than 90 degrees to exact alignment with the applied field, the change may occur only by *rotation of the vectors*, since no

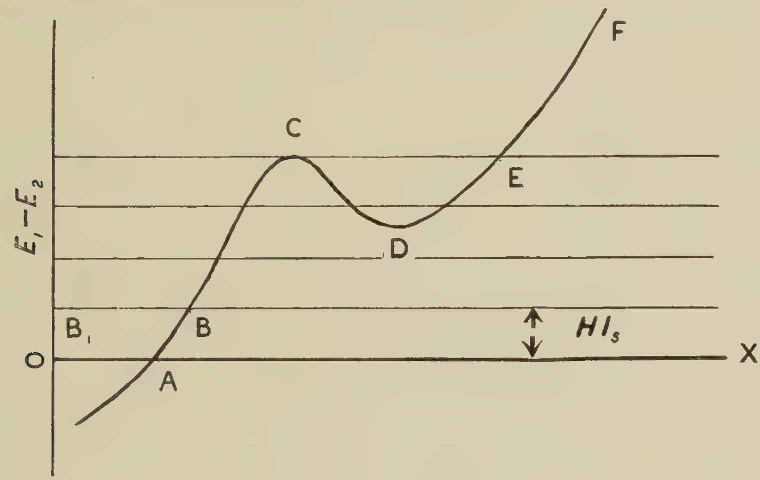


Figure 2. Illustrating energy changes involved in boundary displacements.

further boundary displacements are possible; strong fields must be used and, consequently, $HI_{0,T} \gg \lambda_s F_i$. Hence, when the boundary displacements found on the steep part of an (I, H) curve come to an end and rotations commence, a marked kink appears in the curve. As far as is known these rotations are reversible. It is actually possible to harden a polycrystalline nickel wire by cold working so that it is isotropically and homogeneously strained to such an extent that $E_F \gg E_c$, and all magnetization of the wire occurs by rotations alone. Usually, of course, in polycrystalline materials, boundary displacements and rotations are superimposed and their effects are often difficult to disentangle. This is particularly the case with exceptionally hard materials, like steels, where irreversible processes are evident in all changes in magnetization.

An attempt will now be made to examine how boundary changes and rotations occur in the several portions of a typical hysteresis cycle for a soft material.

Referring to figure 3, we consider first the initial portion Oa of the virgin curve, whose slope or initial susceptibility is determined by the relative numbers of 90- and 180-degree boundary changes. In general, for annealed specimens, the 90-degree changes predominate, and the number of 180-degree boundaries depends on the method by which the specimen was demagnetized, e.g. by reversals with D.C. or by gradually decreasing an A.C. field, prior to the measurement of I and H. We know that the path Oa is magnetically reversible, so that the boundary changes there are most probably reversible too, and we therefore conclude that only in exceptional circumstances is the number of 180-degree boundary

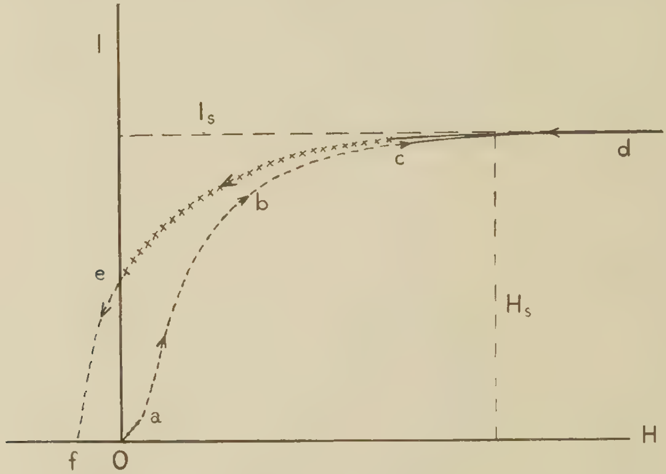


Figure 3. Interpretation of the hysteresis curve.

- - - - - - - - Reversible 90 and few 180 degree boundary displacements.
— — — — Mainly irreversible 180 degree boundary displacements.
x x x x x x x x Mainly reversible 90 and 180 boundary displacements.
———— Reversible rotations.

changes large in this region; indeed, it appears to be large solely in the case of carbonyl iron. Approximate expressions for the initial susceptibility may therefore be deduced.

Thus, for a nickel rod under severe internal stress (as distinct from severe stress caused by external forces which cause all the domain vectors to set perpendicular to the axis of the rod and give an initial susceptibility $k_0 = \frac{1}{3} \,.\, I_s^2/\lambda_s F_i$) k_0 can be shown (Becker and Döring, 1939) to be approximately equal to $\frac{2}{9} I_s^2/\lambda_s F_i$. The latter expression, strange as it may seem, differs but little from that for a substance with F_i very small and for which 90-degree boundary displacements are of overwhelming importance. Assuming, therefore, that the minimum value of F_i may be taken equal to $\lambda_s Y$, the maximum value of the initial permeability could not greatly exceed $1 + 4\pi \,.\, \frac{2}{9} I_s^2/\lambda_s^2 Y$, i.e. about 12 000 e.m.u. per c.c.; this is in good accord with some of the experimental values for pure iron, which could not be explained on a simple rotation picture, but is much lower than the

values obtained with specially treated materials, which are due, presumably, to 180-degree boundary displacements.

Returning to figure 3, the steep part *ab* is mainly due to irreversible 180-degree boundary changes. The bend or " knee " of the curve which is produced by the superposition of huge numbers of the sharp kinks found with single crystals is located at *b*, where the effective field is approximately equal to the coercivity H_c. This knee represents the transition from the region of 180-degree boundary changes to the region *cd* in which nearly all changes in magnetization are due to reversible rotations. If we increase the field beyond the point *d* we leave the region of technical magnetization processes and enter the region, so well explored in magneto-caloric experiments, where further magnetization results from the reorientation of individual electron spins alone.

On decreasing the field the magnetically reversible portion of the path *de* is traversed, and the magnetization thereafter decreases because of reversible 90-degree and, possibly, a few 180-degree boundary displacements, until the point *e* at the intersection of the curve with the I axis is reached. The field is then reversed, and a little beyond *e* the reversible boundary displacements are completed when the curve descends steeply to *f* because of irreversible 180-degree boundary displacements. There is no advantage in discussing the remaining stages of the cycle as all the essential features can be appreciated from the above description.

The intercept I_r on the I axis is the retentivity. In the case of magnetically soft materials it is easy to see, from a consideration of the symmetry of the distribution of the vectors about the direction in which the field was formerly a maximum and the absence of vectors pointing in the reverse direction, that I_r should be equal to $I_s/2$. However, there is usually sufficient magnetic interaction between neighbouring domains to make $I_r > I_s/2$. It should be noted that the coercivity *Of* cannot be greater than the residual induction, $4\pi I_r$.

In the case of heavily strained materials there should be very little difference between the demagnetized state and the state $I = I_r$, the only real difference being that certain of the domain vectors are oppositely directed in the two states. Hence the differential susceptibility k_r at $I = I_r$ should be equal to the initial susceptibility, and this is supported by experiment. Thiessen (1940) has found for a nickel wire in which homogeneous, isotropic, large internal strains have been produced by special cold working, that $k_0 = \frac{1}{3}\frac{I_s^2}{\lambda_s F_i}\left(\frac{2}{3} - \frac{2}{15}\frac{F}{F_i} + \ldots\right)$ when an external stress F is applied.

For lightly strained materials we expect a different relation to hold. In the case of iron, Becker calculates that $k_r/k_0 = 1 - (2/\pi) = 0{\cdot}364$, while for nickel the smaller ratio 0·328 is obtained. In other words, the value of the ratio k_r/k_0 should never fall below 0·3, but, if in a rare case this should happen, as, apparently, it does with carbonyl iron, it is easy to explain it by postulating that 180-degree as well as 90-degree boundary displacements are involved.

I now wish to examine in some detail the information which we may obtain from the heat changes which accompany magnetization processes. Different materials, of course, show very different heat losses, and I remind you of the very wide range of losses exhibited by iron-nickel-cobalt alloys.

I now propose to describe to you some experiments carried out by Dr. J. C. Weston and myself which were recently published in the *Proceedings* of the Physical Society (Bates and Weston, 1941). It is generally accepted that the energy liberated in a complete hysteresis cycle is given by the expression

$$\oint_{H}^{H} H \, . \, dI,$$

a statement which is sometimes known as Warburg's Law. Let us assume that HdI correctly represents the whole energy supplied to a ferromagnetic specimen in any part of a hysteresis cycle and that it is accompanied by an increase dE in internal energy. For the present we will confine attention to specimens which are not subjected to applied mechanical forces, neglecting those provided by the presence of the atmosphere, so that we may write $dE = d(E_c + E_F)$, where E_F refers only to the energy associated with strain of internal origin. The changes are adiabatic, so that we may write

$$HdI = (\partial E/\partial I)_T dI + (\partial E/\partial T)_I dT. \qquad \ldots\ldots(1)$$

Now, HdI can be obtained directly from magnetic measurements, while the second term on the right-hand side of equation (1) represents that part of the change in the internal energy which manifests itself as an increase in temperature, and can be obtained directly from thermal measurements now to be described. In fact, we may write

$$(\partial E/\partial I)_T dI = HdI - dQ_1, \qquad \ldots\ldots(2)$$

and we may therefore determine $(\partial E/\partial I)_T$ at any stage in the magnetization of a ferromagnetic, which means that we can say whether E increases, decreases or remains constant with change in I at any point on an (I, H) curve, equation (2) being independent of whether the change is reversible or not in a thermodynamic sense. Equation (2) brings out the importance of considering thermal changes with reference to changes in I instead of to changes in H. The thermal measurements were made by a new method which has been described in detail in these *Proceedings* this year, so that I need only refer to them briefly here. A number of thermocouple circuits were used; the "hot" junction of each thermocouple was placed in direct contact with the ferromagnetic specimen with the "cold" junction quite close thereto, but thermally insulated therefrom except for conduction along the material of the couple. Each thermocouple circuit was connected in series with its own insulated primary winding, of some 20 turns of low resistance wire wound on a mu-metal spiral core or ring which served as the core of a transformer. The secondary coil of 2 000 turns of low resistance wound on the core was connected to a special form of moving coil galvanometer or fluxmeter. The latter was peculiar in that the magnetic field was provided by an

electromagnet while the ferromagnetism normally found in its moving coil was artificially increased to provide a negative restoring couple, i.e. a couple in opposition to that provided by the suspension, so that an instrument of long period and very high sensitivity resulted.

The temperature-measuring system was calibrated by producing standard adiabatic changes of temperature of the order of 0·001° c., by applying known longitudinal stresses to the specimen. Joule showed that when a tension of F dynes is suddenly applied to a metal rod there occurs an adiabatic fall of temperature ΔT given by

$$\Delta T = \frac{\alpha \, . \, T}{J\rho C} \cdot \frac{F}{A}, \qquad \ldots\ldots(3)$$

where α is the coefficient of linear expansion of the rod, T the absolute temperature J the mechanical equivalent of heat, and ρ, C and A are respectively the density, specific heat and area of cross-section of the rod. We see then that the application of F really causes the absorption of a quantity of energy $J \, . \, \rho \, . \, C\Delta T$ per c.c. of the specimen; this is particularly helpful, since in hysteresis experiments it is the heat liberated or absorbed per c.c. which is measured.

We will now consider a few representative results obtained by Weston and myself, working mainly with nickel, because this metal can be obtained in a very pure state, its magnetostriction is moderately isotropic, and it has been subjected to close theoretical study. In figure 7 of the paper by Bates and Weston are plotted the data recorded for a hard-drawn rod of nickel, 99·98% pure, with less than 0·001% S, less than 0·02% C and a trace only of other impurities. The data were obtained by starting with the nickel in a field of 194 oersteds, decreasing the field to zero and then increasing it to 194 oersteds in the opposite direction. Starting from the left-hand side of the figure, the full-line curve represents the successive values of ΣdQ_1, or Q_1, obtained by adding the several steps of dQ_1, plotted against the appropriate values of I. The broken line represents the $\left(\int HdI, I\right)$ curve calculated from the magnetic measurements made by the ballistic method, and the dotted line represents the $\left(\int HdI - Q_1, I\right)$ curve. In order to save space the additional data which were obtained by reducing the field to zero once more and then increasing it to its maximum value in the original direction have not been plotted. They would merely lie on a similar (Q_1, I) curve which can be obtained by rotating the printed full-line curve about the axis of ordinates while displacing it parallel to that axis until the first point on the left-hand side rests upon the furthermost point on the right-hand side. The experimental values of the retentivity are indicated by I_{rem} or I_r on figure 7 and the remaining figures in the paper.

In figures 15 and 16 of that paper are shown the results obtained with specially annealed nickel of the same composition as that used for figure 7 when maximum fields of 200 and 400 oersteds respectively were used.

The contrast between figures 7 and 15 brings out the profound effects of annealing, while that between figures 15 and 16 indicates the effects of applying fields sufficiently high to magnetize the specimen beyond the regions in which kinks are observed in the (I, H) curves for single crystals of nickel. In all these figures it is seen that when $H=0$ and $I=I_r$, $(\partial E/dI)_T$ is always zero. This means that $(\partial(E_c+E_F)/\partial I)_T=0$ when $I=I_r$. In other words, in these circumstances the changes in crystal energy are exactly compensated by the changes in strain energy. Indeed, for one particular specimen of nickel 99·67 % pure, containing 0·04, 0·02, 0·04, 0·08, 0·01 and 0·14% of C, Si, Cu, Fe, Mn and Mg respectively, it was found that $(\partial E/\partial I)$ was equal to zero for all values of I; but this specimen was exceptional. The lower the coercivity of the nickel specimen the more pronounced was the minimum value of $\left(\int HdI-Q_1\right)$. Moreover, $(\partial E/\partial I)_T$ was negative whenever the magnetization processes were magnetically reversible and positive when they were magnetically irreversible, a statement which also applies to magnetization experiments with virgin material.

For all the (hard-drawn) nickel-iron alloys studied, $\left(\int HdI-Q_1\right)$ was always positive for the magnetization cycles, $(\partial E/\partial I)$ being zero when $I=I_r$ and practically zero over a wide range of I extending from this state. There appeared to be no essential difference other than magnitude between the thermal behaviour of mu-metal and permalloy, which show very small magneto-striction, and 42 % nickel-iron alloy which has a very large volume magneto-striction. The results for a hard-drawn 36·3 % nickel-iron alloy whose constitution approximates to that of invar is shown in figure 25 of the paper.

In all the experiments, Warburg's law was found to hold with an accuracy of better than one per cent for a single cycle. In fact, Bates and Weston used the law to determine the coefficient of linear expansion of invar. They first observed for a complete cycle the sum of the deflections of their temperature-measuring system corresponding to Q_1. This sum was equivalent to the quantity of heat which they calculated from $\oint_{+H}^{+H} HdI$. In other words, the magnetic data were used for calibration purposes. The invar rod was then loaded suddenly and the deflection measuring the heat absorbed was noted, and α was calculated from equation (3), the value so found being $0·411\times10^{-6}$ per °C.

The behaviour of nickel under severe tension is shown in figure 23 of the paper. It is immaterial whether the specimen is originally annealed or hard-drawn. The thermal changes are very small and, consistent with the Becker's theory, there is no sign of cooling anywhere in the cycle. Actually, traces of reversible heating and cooling were found with a load of 8·57 kg. per mm^2, but there was no cooling at 12·2 kg. per mm^2, in agreement with Kirchner's finding that the magneto-striction varied linearly with I^2 only when the load exceeded 10 kg. per mm^2.

In this lecture I have tried to explain to you the importance of the modern

domain conception of magnetization processes. We have seen how it is supported by experiments on the Barkhausen effect, on the magnetic properties of nickel and permalloy under strain, and, finally, by the experiments on heat losses. The theory, in my opinion, has come to stay, and I only hope that you have enjoyed listening to this lecture as much as I have enjoyed giving it.

REFERENCES

Bates, L. F., 1931. *J. Sci. Instrum.* **8**, 376.
Bates, L. F., 1939. *Modern Magnetism* (Cambridge), p. 161.
Bates, L. F. and Weston, J. C., 1941. *Proc. Phys. Soc.* **53**, 5.
Becker, R. and Doring, W., 1939. *Ferromagnetismus* (Berlin), p. 154.
Sixtus, K. J. and Tonks, L., 1931. *Phys. Rev.* **1**, 317.
Thiessen, G., 1940. *Ann. Phys., Lpz.*, **38**, 153.

A REFLECTION METHOD OF MEASURING OPTICAL AND ELECTRICAL CONSTANTS AT ULTRA-HIGH RADIO FREQUENCIES

By Professor L. S. PALMER and G. R. FORRESTER,
The University College of Hull

Received 6 *April* 1941

ABSTRACT. A source of ultra-high-frequency oscillations was coupled to a long antenna AB at a point O. By terminating one end B of the antenna with a reflecting surface of the substance under investigation, standing waves were produced between the reflector at B and the input at O. Resonance conditions along OB were detected by a thermojunction in the portion OA of the antenna and were controlled by altering the length OB. The phase change ϕ produced on reflection was calculated from the difference in the resonant lengths OB with and without the reflecting termination. From the current distribution along OB, the amplitude reduction factor ρ was determined. The electrical and optical constants of the reflecting substance were then calculated from their known relationships to the Fresnel reflection coefficient $\rho e^{j\phi}$.

From the measurements with clay soils with different moisture content it was found that the refractive index increased from 2·6 to 3·6, the dielectric constant increased from about 5 to 11 and the electrical conductivity increased from $1{\cdot}5 \times 10^9$ e.s.u. to $1{\cdot}8 \times 10^9$ e.s.u. as the moisture content of the clay increased. The frequency used was 407 megacycles.

It is believed that the values of the phase change ϕ were measured to within 1 per cent, and that the amplitude reduction factor ρ was measured to within 3 per cent.

§ 1. INTRODUCTION

The object of the experiments described below was to determine some of the electrical and optical properties of various substances such as soils, water, etc., for wireless waves from about 70 to 80 cm. in length. High-frequency methods for measuring these properties usually involve some

modification of Drude's Lecher-wire method (Drude, 1897) or entail the use of a suitably designed condenser in which the substance under investigation is the medium between the plates. Smith-Rose and McPetrie (1934) used the former method for measuring the electrical properties of soils, whilst others—the most recent being Cheng (1940)—used the latter method. A reflection method recently described by McPetrie (1934) is based on the early theoretical work of Fresnel. Fresnel has related the refractive index, the extinction coefficient, the dielectric constant and the electrical conductivity of a substance to its reflection coefficient by the well known conceptions of a complex refractive index and a complex dielectric constant. The problem thus resolves itself into a measurement of the reflection coefficient, which, in turn, can be calculated from the absorption and phase change produced when incident electromagnetic waves are reflected from the substance under investigation. The reflection method devised by McPetrie involved the projection of ultra-short wireless waves vertically downwards from an elevated transmitter. After normal reflection from the ground, the incident and reflected waves combined to give a resultant wave, the amplitude of which was a maximum at such distances from the ground that the two component waves differed in phase by an integral number of wave-lengths. The positions of these maximum fields were detected by the occurrence of current maxima in a receiving aerial suspended between the transmitter and the ground. Such maxima were found to be situated $\lambda/2$ cm. apart (λ being the wave-length), indicating a phase change of $\frac{4\pi}{\lambda}\left(\frac{\lambda}{2}\right)$ or 2π as required for successive reinforcements between the incident and reflected waves. The displacement of the peaks from their positions when the reflector was a mat of copper gauze and when it was a given kind of soil enabled the phase change ϕ (relative to copper) to be calculated from the formula* $\phi = 4\pi x/\lambda$, where x is the displacement.

McPetrie calculated the absorption factor ρ from the reduction in amplitude suffered by the wave after reflection. From these data the *Reflection coefficient* R can be determined from the usual equation:

$$R = \rho e^{j\phi}.$$

The method used in the present investigation is a development of work by Palmer and Gillard (1938) on the current distribution in a long transmitting antenna. It was found that the current in that part of a long antenna which was (say) above the point where the power was fed in was a maximum or a minimum depending on whether the impedance of the "effective length" of that portion of the antenna *below* the feeding point was a maximum or a minimum respectively. By "effective length" l_0 (figure 1) is meant the geometrical length l of the lower portion of the antenna plus the equivalent length y of the

* McPetrie used the formula $\phi = 2\pi x/\lambda$, which necessitates a correction to his values for the dielectric constant and conductivity. The former need to be reduced, but it happens that the latter are only slightly affected.

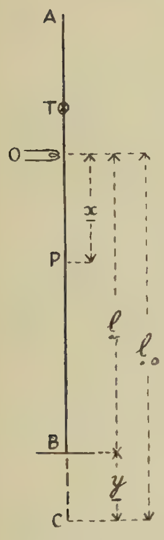


Figure 1.

reflecting termination. This equivalent length of the reflecting termination is defined by the equation $y=\lambda\phi/4\pi$, where ϕ is the phase change between the incident and reflected waves. Thus the difference (l_0-l) determines the phase change produced on reflection at the end of the antenna. By terminating the antenna with a metal sheet or by sticking it into the earth's surface or into a trough of liquid, the phase changes resulting from these several substances may be measured with considerable ease and precision. The method has the advantage that no instrument is incorporated in that portion of the antenna under measurement, the current being measured by an instrument included in the upper portion of the antenna, the length of which is immaterial.

To obtain the absorption coefficient ρ, current-distribution measurements were made along that portion of the antenna near to the reflector. By taking current-amplitude ratios close to the reflector, changes in amplitude due to losses along the antenna itself were avoided.

The elementary theory of the method is given in the following section.

§ 2. THEORETICAL CONSIDERATIONS

In figure 1 let l be the geometrical length of that portion OB of the antenna AB which is below the energy input point O, and let BC be the equivalent length y due to a phase change ϕ after reflection at B. The phase change is equal to the distance travelled by the wave in going from B to C *and back*; that is,

$$\phi=4\pi y/\lambda=4\pi(l_0-l)/\lambda. \qquad \ldots\ldots(1)$$

If the total or resultant current at any point P a distance x cm. from O is I_T, then

$$I_T=I_i\cos\{\omega t-2\pi x/\lambda-(2n+1)\pi\}+I_r\cos\left[\omega t-\frac{2\pi}{\lambda}\{l+2y+(l-x)\}\right],$$

where I_i and I_r are the amplitudes of the incident and reflected waves respectively. Hence

$$I_T=\left(I_i^2+I_r^2+2I_iI_r\cos\frac{4\pi}{\lambda}\{l+y-x-(2n+1)\,.\,\lambda/4\}\right)^{1/2}\times$$
$$\cos\left[\{\omega t-\frac{2\pi x}{\lambda}-(2n+1)\,.\,\pi\}-\alpha\right], \qquad \ldots\ldots(2)$$

where α is the phase lag between the incident and resultant currents.

When the current at any point T (figure 1) in the upper portion of the antenna is a minimum, the length of the lower portion is such that the total reactance of the length l together with the additional length y is zero, and I_T is a maximum and α is zero. It follows from (2) that current maxima will occur for values of x given by

$$x=(l+y)-(2n+1)\lambda/4, \qquad \ldots\ldots(3)$$

that is, at points situated an odd number of quarter-wave-lengths measured from C, a distance of y cm. *beyond* the lower end B of the antenna.

The values of the maximum and minimum currents will be $I_T\binom{\max}{\min} = I_i \pm I_r$.

The modulus $|R|$ of the complex reflection coefficient is equal to the ratio $I_r/I_i(=\rho)$; hence

$$|R| = \frac{I_{T(\max)} - I_{T(\min)}}{I_{T(\max)} + I_{T(\min)}} = \rho. \qquad \ldots\ldots(4)$$

If these values of I_T be measured as closely as possible to the reflecting termination at B, then their ratios will be practically independent of the aerial attenuation. For this reason the resistance and leakance of the antenna have not been considered in the foregoing elementary treatment.

With R in the form $\rho e^{j\phi}$ or $\rho(\cos\phi + j\sin\phi)$, we have (following Fresnel)

$$\eta' = (\eta - jk) \quad \text{and} \quad \kappa' = (\kappa - j\,.\,2\sigma/f),$$

where η' and κ' are the complex refractive index and dielectric constant respectively, η and κ the ordinary refractive index and dielectric constant respectively, k the extinction coefficient and σ/f the ratio of the electrical conductivity to the wave frequency.

Hence
$$\eta' = \frac{1 - \rho^2 - 2j\rho\sin\phi}{1 + \rho^2 + 2\rho\cos\phi},$$

leading to
$$\eta = (1-\rho^2)/(1+\rho^2+2\rho\cos\phi), \qquad \ldots\ldots(5)$$

$$\kappa = \frac{(1-\rho^2)^2 - 4\rho^2\sin^2\phi}{(1+\rho^2+2\rho\cos\phi)^2} \qquad \ldots\ldots(6)$$

and
$$\sigma = \frac{2f\rho(1-\rho^2)\sin\phi}{(1+\rho^2+2\rho\cos\phi)^2}. \qquad \ldots\ldots(7)$$

Instead of calculating the values of κ and σ from the above formulae, they may be read directly from the ingenious graphs published by McPetrie (1934), in which his $K = \rho\cos\phi$, $K' = -\rho\sin\phi$ and $\phi = \tan^{-1}(-K'/K)$.

§ 3. EXPERIMENTAL DETAILS

The oscillator used in the present experiments was that described by Palmer and Gillard (1938) and the wave-length generated was 73·7 cm. The oscillatory circuit was loosely coupled to a straight vertical antenna which was constructed from sections of 5/16-inch diameter brass rods. The sections were of different lengths and could be screwed together so that the length OB (figure 1) could be varied in steps of 1 mm. or more up to a total length of several wave-lengths. The introduction of washers 1 mm. thick was necessary in order to make accurate determination of ϕ. A rod sliding in a closely fitting brass tube was also used for fine adjustment.

For determining the critical values of l, the current at T (figure 1) was measured

by a vacuum thermojunction connected by suitably screened horizontal leads to a microammeter.

Because of the flow of current past O to compensate for unavoidable resistance losses, the reading of the microammeter was never zero. The minimum value was obtained when the lengths AO and OB were adjusted so that the impedance of the former was a maximum and the impedance of the latter a minimum (Palmer and Gillard, 1938, equations (3) and (4)). Furthermore, it was found that the adjustment of OB or l was most sensitive when the length AT was about a quarter of a wave-length. This ensured that the thermojunction at T was situated at a current antinode.

These adjustments are not essential, and theoretically AT and TO can be any lengths whatsoever, and the current at T should always be zero when l is critically adjusted.

Measurements were first made of three or four critical lengths l_{air} for an open-ended antenna (see figure 4). Similar measurements were then taken with different terminations at B. In particular, circular copper discs were screwed on the end of the antenna, or the antenna was lowered into large trays of different samples of soil. Other experiments were made in the open field where the area of the earth-termination was unlimited. Experiments are in progress with various liquids.

It was found in practice that the critical values of l varied appreciably with the size of the terminal plate or tray of soil. Consequently, when an unlimited area could not be procured, as in the the case of copper, it was necessary to obtain a (y, radius) graph (table 1 and figure 2) from which the maximum value of y for a copper disc of infinite radius could be determined. The points on figure 2 were obtained by using terminal discs varying in radius from 1 to 16 cm.

Other factors remaining constant, the capacity of the antenna terminating plate depends directly on its radius. Consequently, from *a priori* considerations it might be anticipated that the variation of the capacity effect with the disc radius would get less and less as the radius approaches infinity. In other words, the effect on the phase change of increasing the radius r of the plate should get less and less as the plate gets larger and larger and as the actual phase change ϕ approaches the maximum value ϕ_0 for a plate of infinite radius.

Hence $d\phi/dr = K(\phi_0 - \phi)$, where K is some constant, i.e.

$$dy/dr = K(y_0 - y), \qquad \text{.......(8)}$$

where $y_0 = \lambda\phi_0/4\pi$ and $y = (l_0 - l_{\text{copper}})$ of figure 2.

Equation (8) has not been established rigorously, but has been interpolated from the experimental data and shown to have a reasonable physical basis. It therefore serves our present purpose in that it enables the required correction to be obtained from the experimental measurements. Thus y_0 (and therefore ϕ_0) for a plate of infinite radius may be determined directly from an intercept of the

(dy/dr, y) graph. The experimental values of r and y and the calculated values of dy/dr are given in table 1.

Table 1

r (cm.)	0	1	2·1	4	6	8	10	12	14	16
y (cm.)	0	3·6	6·7	11·1	13·8	15·1	16·0	16·1	16·2	16·3
dy/dr	3·6	2·81	2·3	1·35	0·65	0·45	0·05*	0·05*	0·05*	

* These figures are less reliable.

From the resulting graph (figure 3) it follows that $y_0 = 18{\cdot}5$ cm., which is $\lambda/4$ to within the errors of the experiments. Thus the values of l_{copper} must be

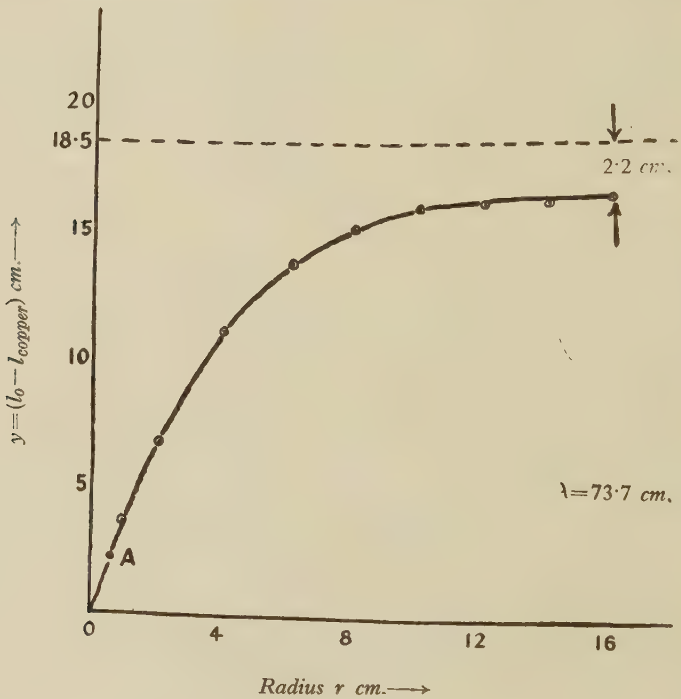


Figure 2.

corrected by an amount (16·3–18·5) or $-2{\cdot}2$ cm. in order to take account of the limited extent of the 16 cm.-radius terminal reflecting plate.

By integrating equation (8) and introducing the boundary condition $y=0$ when $r=0$, we get

$$y = y_0(1 - e^{-Kr}). \qquad \ldots\ldots(9)$$

This is an equation for the graph in figure 2, the asymptote being

$$(l_0 - l_{\text{copper}}) = y = y_0 = 18{:}5 \text{ cm.} \simeq \lambda/4.$$

No correction was necessary for l_{soil} because the measurements given in figure 4 were taken in a field the dimensions of which were practically unlimited compared with the wave-length.

In the case of l_{air} the error was in the opposite sense, in that the area of the end of the 5/16-inch rod was not infinitely small but had an appreciable capacity effect which made the experimental value of l_{air} too short. This error was overcome by tapering the end of the rod to a fine point of negligible area. With an antenna tapered in this way the measured values of l_{air} (given in figure 4 and in table 2 below) were greater by 2·25 cm. than the critical lengths obtained with the untapered rod. This correction may also be obtained directly by reference to point A in figure 2.

The measurements of l and the calculations of ϕ are given and discussed in the following section.

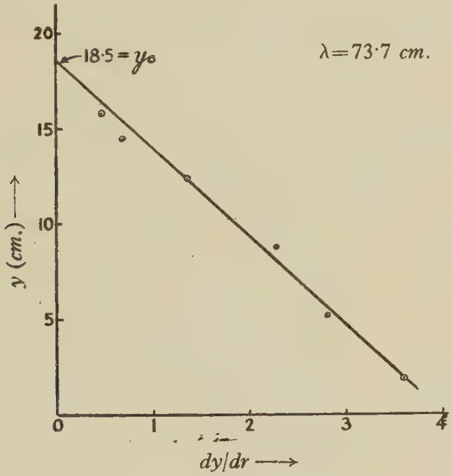


Figure 3.

In order to calculate ρ, current-distribution curves were plotted from data obtained by moving a second thermojunction from B towards O in the lower part of the antenna. The thermojunction was inserted in a short section of brass rod which could be incorporated as a section of the antenna at any chosen position by suitably adjusting the length of the screwed sections above and below the junction. For these measurements it is desirable, but not necessary, to know the effects on l of introducing the thermojunction. This was readily obtained in equivalent cm. of antenna rod by noting the change in l necessary to maintain constant minimum current at T with and without the thermojunction in OB. It was found that the junction, leads and microammeter were equivalent to about 9·5 cm. of antenna rod at the frequency of 407 megacycles used in this work. Although this correction does not affect the measurements of the ratio $I_{T\,(min)}/I_{T\,(max)}$, it does enable current-distribution curves to be shifted along

the axis of abscissae to their proper relative positions, which would otherwise be displaced, due to the shortening effect of the thermojunction.

The value of ρ_{copper} will be too small, owing to the limited area of the reflecting plate, whilst the value of ρ_{air} will be too big, due to spurious reflections from the surroundings. The former error can be eliminated by using a large copper sheet for the reflector, or can be corrected by obtaining a (ρ, r) graph similar to the (y, r) graph of figure 2. In the case of ρ_{air} it was not possible to eliminate stray reflections even with a horizontal antenna. But as neither ρ_{copper} nor ρ_{air} is required for the attainment of the present objectives, the experiments with the open-ended antenna and with the copper plate termination were only pursued as far as was necessary to get reliable values of the critical resonant lengths, with which the equivalent lengths for other reflectors have to be compared.

These difficulties in measuring ρ do not arise in the case of the clay soils, for which an unlimited area was available and for which the ground reflections are an essential factor in the measurements.

§ 4. RESULTS AND DISCUSSION

Figure 4 depicts to scale the values of l for air, clay, wet clay and copper antenna terminations. The actual curve represents a hypothetical incident wave. It is significant that $l_{air} > l_{clay} > l_{copper}$ and hence $\phi_{air} < \phi_{clay} < \phi_{copper}$, leading to the conclusion that all phase changes on reflection are less than 180°, as is required for a positive value of the electrical conductivity [equation (7)]. The sensitivity of the current variation at T (figure 1) was such that, in the case of air and copper, l could be adjusted to about 2 mm. and in the case of the clay soils to about 3 mm. Since $4\pi/\lambda$ is 9·8 with the wavelength of 73·7 cm., the accuracy to which ϕ could be measured was about 2° in the case of copper and about 3° in the case of soils. These values are relative to air, for which zero phase change was assumed, that is, l_0 (figure 1) is experimentally determined by measuring l_{air} with the tapered antenna.

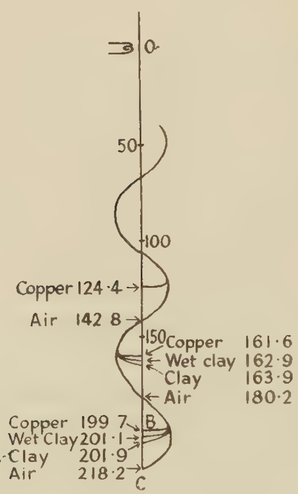


Figure 4.

The values for the copper reflector in figure 4 have been corrected by subtracting 2·2 cm. from the actual measured antenna lengths in accordance with figures 2 and 3. Thus the values of y for copper are

$$\left.\begin{array}{l} 142{\cdot}8 - 124{\cdot}4 = 18{\cdot}4 \text{ cm.} \\ 180{\cdot}2 - 161{\cdot}6 = 18{\cdot}6 \text{ cm.} \\ 218{\cdot}2 - 199{\cdot}7 = 18{\cdot}5 \text{ cm.} \end{array}\right\} \text{giving a mean of } 18{\cdot}5 \text{ cm.}$$

Table 2

Substance	Reference	λ (cm.)	l (cm.) (figure 4)	$y=(l_0-l)$ (cm.)*	Mean y (cm.)	ϕ (°)= $720\,y/\lambda$	ρ	Refractive index η	Dielectric constant κ	Electrical conductivity σ (e.s.u.)
Clay	Present result	73·7	163·9 201·9	16·3 16·3	16·3	160	0·55	2·6	4·8	$1{\cdot}5\times10^9$
Dry soil	McPetrie (1934)	150	—	—	5†	156‡	0·52 to 0·55	2·9 § to 2·35	3·5 to 3·3	$6{\cdot}0\times10^8$ to $7{\cdot}2\times10^8$
Wet clay	Present result	73·7	162·9 201·1	17·3 17·1	17·2	169	0·6	3·6	10·8	$1{\cdot}8\times10^9$
Damp soil	McPetrie (1934)	46	—	—	2	149‡	0·6	1·94 §	0·3	$2{\cdot}3\times10^9$

* $l_0=l_{air}=180{\cdot}2$ cm. and 218·2 cm. as figure 4.

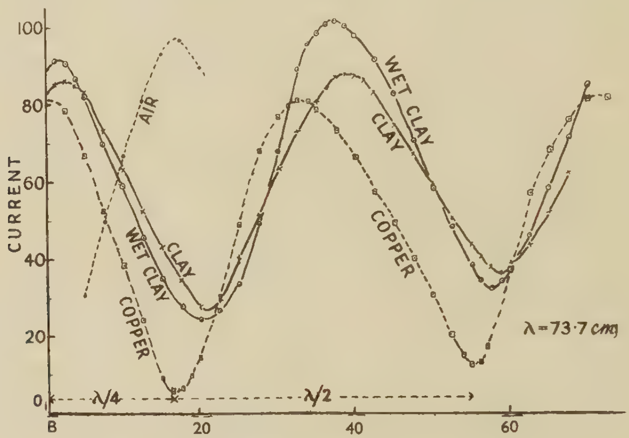
† Measurements referred to copper and corrected by 2 cm. for copper zero-error.

‡ ϕ values subtracted from 180° in order to compare with ϕ_{air}.

§ Calculated from data by McPetrie (1934).

Thus $$\phi_{\text{copper}} = \frac{4\pi \times 18{\cdot}5}{\lambda} = \frac{720 \times 18{\cdot}5}{73{\cdot}7} = 180^\circ{\cdot}6.$$

In order to calculate the optical and electrical constants of copper it is necessary to measure both ϕ and ρ with much greater accuracy than the present refinements of this method permit. Consequently, only the data for clays are included in table 2 above. It may be noted, however, that, when l_{copper} is corrected for the limited area of the reflecting plate, the corresponding values of l_{air} and l_{copper} differ by $\lambda/4$ to within 0·5 per cent. Therefore either an open-ended antenna or one with a copper reflector might be used as the standard from which other phase differences could be compared with little loss in accuracy. But because of the inability to obtain a very large sheet of copper, the critical resonant lengths for air have been used as the standards of comparison in this work.



Distance from terminal reflector B (cm.)=(l−x) cm.

Figure 5.

For convenience, the curves have been shifted relative to each other, as indicated above.

The critical lengths for clay soils and the resulting values of ϕ calculated from equation (1) are collected together in table 2.

Typical current-distribution curves are shown in figure 5. The ratios $I_{T\,(\text{min})}/I_{T\,(\text{max})}$ for clays have been determined from the nodes and antinodes nearest to the origin. The resulting values of ρ (namely, 0·55 for clay and 0·6 for wet clay) are given in the table. It is difficult to estimate the accuracy to which ρ has been determined, but it is probably not more than about 3 per cent.

From these data the refractive index, the dielectric constant and the electrical conductivity of clay soils with different degrees of wetness have been calculated from formulae (5), (6) and (7), the frequency being 407 megacycles. For comparison, McPetrie's determinations (recalculated in accordance with the footnote on page 480) are included.

§ 5. CONCLUSION

It is difficult to compare the various results given in table 2 because of the different natures of the soils, the different degrees of wetness and the different frequencies employed. But, with the exceptions of the low values of η and κ for damp soil (McPetrie, 1934), they are all of comparable magnitudes respectively. It would appear that the dielectric constants are, on the whole, rather lower than the value of 10 which was assumed by McPetrie. On the other hand, the values of the electrical conductivities, as measured in the present work, lie between the extreme values calculated from the data of McPetrie. The mean value is about $1{\cdot}5 \times 10^9$ e.s.u. Because the present measurements were made on one wavelength—73·7 cm.—it is possible to study the effect of moisture on the various constants. As may be anticipated from theoretical considerations, all three constants increase with increase in the moisture content of the soil—the refractive index from 2·6 to 3·6, the dielectric constant from about 5 to 11 and the electrical conductivity from $1{\cdot}5 \times 10^9$ e.s.u. to $1{\cdot}8 \times 10^9$ e.s.u. These changes cannot be related quantitatively with the moisture content of the soil, because no specific analyses were undertaken in the experiments described above. Some controlled experiments on wet clays are now in progress from which it is hoped that a definite relationship between the electrical and optical constants and moisture content will be forthcoming.

REFERENCES

Cheng, 1940. *Phil. Mag.* **30**, 305.
Drude, 1897. *Ann. Phys., Lpz.*, **61**, 466.
McPetrie, 1934. *Proc. Phys. Soc.* **46**, 637.
Palmer and Gillard, 1938. *J. Instn. Elect. Engrs*, **83**, 415.
Smith-Rose and McPetrie, 1934. *Proc. Phys. Soc.* **46**, 649.

ON ULTRA-HIGH-FREQUENCY OSCILLATIONS GENERATED BY MEANS OF A DEMOUNTABLE THERMIONIC TUBE HAVING ELECTRODES OF PLANE FORM, by W. A. LEYSHON

DISCUSSION*

Mr. W. E. BENHAM. The results in this paper appear to be of considerable interest, particularly those illustrated by figure 11. A few questions arise:—

(1) On page 143 it is stated that the ($\log i_g$, $\log V_g$) curves were all straight lines with a slope of 3/2 for space-charge limited currents. How exactly is this true? Could a small discrepancy be traced due to initial velocities?

(2) In figure 9, grid volts are taken as $V_g - \frac{1}{2}V_f$. Was this done for the other figures?

(3) In connection with the experiment described on page 150, was it possible to obtain good results with G_2 considerably below K_1 in potential? This arrangement is more suitable in the brake field triode than to take $V=0$ at the plate. Was there any means of measuring oscillation efficiency?

(4) On page 152, the range of oscillation for a diode is given as $0{\cdot}97T<t<1{\cdot}31T$. This should more exactly be $T<t<1{\cdot}43T$ (opt. $1{\cdot}21T$). However, my calculations, allowing for grid transparency (provisional, but based in part on previous published work), show that for the first oscillation range

$$1{\cdot}19T<t_1<1{\cdot}67T \text{ (opt. } 1{\cdot}43T)$$

should hold for the space next to the filament, and apparently

$$0{\cdot}37T<t_2<1{\cdot}00T \text{ (opt. } 0{\cdot}66T)$$

should hold for the space between the grids.

The latter value is calculated ignoring space charge (justified by high electron speeds). Both sets of figures are obtained assuming that the amplification constant μ is very large.

The use of a plane instead of a cylindrical geometry removes the limitation that the grid-plate space alone can contribute negative conductance. It should thus be possible to arrange matters so that both spaces are contributing maximum negative conductance. This will be the case if

$$t_1=1{\cdot}43T \quad \text{and} \quad t_2=0{\cdot}66T.$$

Then will $t_1=2{\cdot}16\ t_2$.

If, further, the current in the KG space is just space-charge-limited and that in the G_1G_2 space is insufficient to affect appreciably the potential distribution there,

$$t_1/t_2=3d_1/d_2\ ;$$

and, therefore,

$$d_1/d_2=0{\cdot}72.$$

* See this volume, part 2, p. 141.

In the above, T = periodic time of generated oscillation,
t_1 = transit time from K to G, $KG = d_1$,
t_2 = transit time from G_1 to G_2, $G_1G_2 = d_2$.

It is understood that the above estimate is subject to further calculation.

AUTHOR. I should like to thank Mr. Benham for his interest in the results given in the paper, and for giving the ranges of oscillation his theory predicts for my tube.

Taking in turn the points he has raised—

(1) The characteristic ($\log i_g$, $\log V_g$) curves were drawn in the first instance as a test of the degree of vacuum in the experimental tube; the results were probably not accurate enough to detect a small discrepancy due to initial velocities; the mean values of the slopes of the graphs, however, differed in no single instance by more than 1·4 per cent from 3/2.

(2) The value ($V_g - \frac{1}{2}V_f$) was used in finding $\lambda\sqrt{V_g}$ ranges and in the ($\log i_g$, $\log V_g$) graphs. The characteristic curves in figures 4, 6, 8 and 10 (*a*) were drawn taking V_g as actually measured (from the negative end of the filament).

(3) My records show that system 1 would oscillate with G_2 at -18 v. with respect to the negative end of the emitting filament K_1, G_1 being at $+160$ v.

In general, however, I found in using the various systems as triodes that decreasing V_{g2} from zero (K_1 emitting) diminished the strength of oscillation. I had no absolute method of measuring oscillation efficiency, but relative efficiency is indicated by the value of $\Delta i_g/i_0$, where Δi_g = maximum change of grid current from the steady value i_0 corresponding to optimum conditions of oscillation.

(4) The diode range was inaccurately quoted in the paper, so that the frequency interval, given as 1·35 on page 153, should read 1·43. The new range of oscillation for the KG spaces predicted by Mr. Benham corresponds to the slightly smaller value 1·40.

The observed frequency intervals given in table 2 (page 153) are of the order predicted for this space; they do not in any single instance approach the value 2·7 predicted for the G_1G_2 space when it alone provides the negative conductance, but this would hardly be expected, since both spaces probably contribute to the resultant conductance.

Mr. Benham was correct in assuming that μ was very large for the systems used.

The new ranges of oscillation given by Mr. Benham for my tube, and his prediction of an optimum value of d_1/d_2, have led me to write a short note on oscillation generation in it, which I hope will be published in these *Proceedings* in due course.

REFERENCES

BENHAM, W. E., 1931. *Phil. Mag.* **9**, 457.
BENHAM, W. E., 1938. *Proc. Inst. Rad. Engrs*, **26**, 1093.

REVIEWS OF BOOKS

Test for Night Vision, designed by W. D. WRIGHT. (London: Sir Isaac Pitman and Sons, Ltd., 1941.) 25*s*.

Dr. Wright has succeeded in his object of providing a simple apparatus for testing night vision, which can be operated by persons with no special training in photometry or sight-testing. Nine cards are supplied, each printed with a broken circle of $4\frac{1}{4}$ inches diameter on a black ground, the lightness of the broken circle diminishing in steps from white (card No. 1) to a dark grey (card No. 9). The cards can be exposed in turn in the simple wooden framework supplied, which also carries the light source, a small disc of radium luminous compound held in front of the card at a distance of 7 in. The test must be carried out in a completely darkened room, and consists in determining the card of highest number for which the dark-adapted observer can still perceive the position of the gap in the broken circle when viewing the card at a distance of about a foot.

It was a good idea to use the slowly decaying radium compound as a standard light source, the " life "—about eighteen months—not being so short as to make renewal irksome. The production of printed cards of given and unchanging reflection factors is difficult, but the requirements here are not severe, and the publishers should be able to maintain the necessary constancy. The brief but clearly written instructions explain that about 49 per cent of observers can see card No. 7, 23 per cent can see card No. 8, and so on. This information, and a table showing the relative amounts of light required by observers who can just see the different cards, provide a basis for interpreting the results of the test. The approximate value of the illumination of the cards, or better, since the luminescent light is green, the equivalent (scotopic) white illumination, might well have been stated.

Especially in these days, many are interested in grading personnel with respect to night vision, and to all these Dr. Wright's test may be recommended. W. S. S.

Why Smash Atoms?, by A. K. SOLOMON. Pp. xii + 174. (Cambridge, Massachusetts: Harvard University Press; London: Sir Humphrey Milford at Oxford University Press, 1940.) 14*s*. net.

Complete success in the writing of a popular or semi-popular scientific book is most difficult to achieve. However, in the present instance the author has made a most gallant and spirited attempt, and even if he has not entirely avoided the statement of half-truths, the *bêtes noires* of any such popular presentation, he has at least given some very clever analogies, and has included a most splendid collection of photographs as illustrations to the text. Some of the most pleasing aspects of this book are, however, the little biographical details which the author has included when referring to scientists of note. The physics of the book is a trifle disappointing ; in many instances the author seems to have become a little incoherent and has conveyed a meaning a shade different from that which he intended. W. B. M.

Practical Solution of Torsional Vibration Problems, Vol. I, by W. Ker Wilson. Second Edition. Pp. xx + 731. (London: Chapman and Hall, Ltd., 1940.) 42*s*. net.

The need arises from time to time for painstaking authors who render valuable services by compiling and editing full information regarding a particular niche of knowledge. Dr. Ker Wilson is such an author, and he is to be congratulated on the thoroughness with which he has accomplished his task. The theme of his book is practical and of no inconsiderable importance. His six chapters, comprised within some seven hundred pages, deal with the following items :—I, Torsional vibration ; II, Natural frequency calculations ; III, Equivalent oscillating systems ; IV, Flexible couplings ; V, Geared systems ; VI, Determination of stresses due to torsional vibration. In addition, there is an appendix on the effective inertia method of torsional-vibration analysis, with special reference to aero-engine/airscrew systems; also numerous examples taken variously from marine, electrical, aeronautical and automobile engineering.

The author has attempted to deal with difficult and complicated problems by elementary methods, and on the whole he may be said to have done this successfully. But the avoidance of "mathematical" methods does not necessarily result in simplification. For instance, in fairly common systems with the crank masses equally arranged, concise expressions can be and have been found for frequency equations on a trigonometrical basis, in contrast to the elaborate algebraic processes which go on growing with increasing number of crankshaft throws. Thus the example, given on p. 59, of a generator direct-coupled to a six-throw in-line engine can be done in a few lines as an elementary exercise in trigonometry, and the amount of labour involved is almost independent of the number of throws. The computation thus saved is considerable in cases where the engine-frequency curves are to be coupled with the airscrew-frequency curves of an engine/airscrew system. In connection with this last mentioned problem, the reviewer finds it difficult to understand the author's statement on p. 693 :—

> "The plotting of the tuning curve for the engine or basic system presents no serious difficulties, however many masses and shafts compose the engine system and whether the engine is geared or not. The real practical difficulty arises when attempting to plot the effective inertia curve for the airscrew, since the mass-elasticity characteristics of the airscrew blades cannot be correctly represented by a series of concentrated masses connected by massless shafts, and the problem is so complex that no complete mathematical analysis has yet been made."

Regarding this it is as well to remember that it has become customary for crankshaft systems to be more or less arbitrarily represented by a series of pulleys connected by massless shafts of empirical flexibility in order to obtain a tractable basis of solution. It is realized that some such artifice must be adopted owing to the limitations of the analytical machine and the imperfectly understood or unknown conditions which arise in the problem of finding the frequencies and modes of vibration of other than an idealized system. The order of accuracy achieved in regarding the airscrew as made up of a series of discrete masses connected elastically by massless members is certainly on a par with that which can be expected from the usual treatment of the engine system. Indeed, the method generally adopted for the engine system, which method is given in the book under review, cannot be regarded as more than a rough approximation, although it does often give results which agree reasonably well with practical observation, at any rate for the fundamental frequency, and in some instances the first overtone.

The book deals with torsional vibration in a most exhaustive fashion, and, having regard to the assumptions made, is sound both in general and in detail. There is no

question that it will rank as a major contribution to the literature of the subject, useful to practical engineers and research workers alike.

The bibliography is somewhat unbalanced. For instance, there is no mention of Morris's *Strength of shafts in vibration* nor any of the numerous contributions he has made to the subject matter of the book under review, yet there is mention of two of his early papers on the whirling of crankshafts which have no bearing on the problem of torsional vibration.

It is rather strange to find no mention of Rayleigh in a treatise on vibration, particularly as Rayleigh's influence coefficients play so important a part in the solution of vibration problems (see, for instance, *Rayleigh's Principle*, by Temple and Bickley).

Also, a mention might have been made of the pioneer effort of Chree, Sankey and Millington, who appear to have been the very first writers in this country to draw attention to the possibility of resonant torsional vibration in power-driven systems.

The Identification of Molecular Spectra, by R. W. B. Pearse, D.Sc. and A. G. Gaydon, Ph.D. Pp. viii+221. (London: Chapman and Hall, Ltd., 1941.) 42*s*. net.

None of the books that have hitherto appeared on molecular spectra has been deliberately planned, as this one is, as a guide to the practical spectroscopist in the often troublesome work of identifying the bands in his spectrograms. It is true that some of them, written by practising spectroscopists, include many observational data in tables, in diagrams and in reproduced spectra ; but the manner of presentation of such material in other books is adapted to the theme of the text, namely, the analysis of band systems and individual bands of various types in accordance with the quantum theory of the structure of molecular spectra ; emphasis is naturally laid on the derivation and tabulation of numerical constants for the electronic states and the vibrational and rotational levels of the molecules. To meet other needs of the laboratory spectroscopist, especially the identification of bands in a composite spectrum, demands an entirely different plan, such as that of the book now before us, which is, in a very real sense, supplementary to works of the kind just referred to. This Society's *Report on Band Spectra of Diatomic Molecules*, for example, may now be said to have in Pearse and Gaydon's tables a particularly appropriate supplement, which is not less welcome, and far more extensive, than if it had appeared with the *Report* nine years ago.

Copies of this book will quickly find their way into spectroscopic laboratories throughout the civilized world as a necessary item of their equipment, and will be well thumbed by many workers who wish to avoid the old troubles and pitfalls. For example, an investigator, not always an entirely inexperienced one, may observe in the spectrogram of a discharge in a carefully prepared and cleaned tube containing highly purified materials some ultra-violet bands which he cannot immediately recognize and which may belong to the very system he is seeking for a certain molecule MX ; he measures the band-heads and makes a more or less complete and correct vibrational analysis. With care he may now succeed in identifying the system, either by a search through original papers or by a calculation of the positions of band-heads from a vibrational formula with coefficients from a table such as Appendix II of this Society's *Report*. Failing that, he will do well to supply data or prints to his colleagues or others who may be familiar with the bands. The bands will, not infrequently, turn out to be part of the

CO Third Positive, the CO Fourth Positive or the CS ultra-violet system, or a few of Schumann's O_2 absorption bands due to air in the optical path outside and inside the spectrograph. While glad to have avoided a wrong assignment, he will be surprised and disappointed to know that his tube contains impurities in such easily appreciable quantities. Many a less cautious and less experienced observer has taken no such precautions, but rushed into print with his scanty data and a highly speculative discussion of the energy levels and behaviour of his particular emitter MX ; thus, the CS molecule, to take but one example, has appeared in the literature several times erroneously attributed to S_2 or other emitters.

Proper use of Pearse and Gaydon's book will obviate all this.

The book consists mainly of two sections : (i) a table of data for 1532 persistent band-heads arranged in order of wave-length from 10603 A. to 2006 A.—corresponding to, say, Kayser's *Tabelle der Hauptlinien der Linienspektra aller Elemente* ; and (ii), the much larger section, a set of lists of data for separate band systems, arranged in alphabetical order, of the neutral and ionized molecules (254 diatomic and about 34 tri- and polyatomic) to which the systems are attributed. From a comparison of the wave-length and degradation of an unrecognized strong band with the data listed in (i) we obtain a clue to the identity of the system and its emitter ; then, by reference to the particulars for that system in (ii), we obtain a check from the presence or absence of other bands of the system. This process of double reference is now repeated for bands of other systems of the same molecule and of any other molecule having an element in common with it ; e.g., having identified a system of C_2 and one of N_2, we look for other systems of C_2 and N_2 and for systems of CN, CO and NO (since air or O_2 may not be entirely absent from the source).

The importance of atomic lines in this work is emphasized, and a table of about 400 persistent lines of elements in alphabetical order of their chemical symbols is provided. In view of the utility and rapidity of direct comparison as a means of identification, six plates of well chosen and beautifully reproduced band spectrograms are included. In addition, there are two plates of comparison spectra (Fe, Cu, Hg, Ne) for use in wave-length measurement. The work is completed with a ten-page section containing valuable practical hints on matters which are often troublesome to the beginner and seldom treated in text-books, and a six-page index of authors and subjects.

That the planning of the tables and the method of using them, as briefly outlined above, are admirably suited to the purpose of the work has been proved in advance by an extensive and successful trial in the laboratory, where the authors' colleagues (including the reviewer and his collaborator) were privileged to use the MS as the plan of the work was being evolved in the last year or two before the war. In a work of this magnitude it cannot be expected that perfection will have been attained at once ; indeed, in the present perusal of it the reviewer has detected a few points where minor typographical improvements might be made in the second edition, which, it may confidently be hoped, the book will reach in spite of the present difficult conditions. Notes of these have been passed on to the authors, who, no doubt, will welcome similar notes from other users of the book.

It is most fitting that this excellent work, which every spectroscopic laboratory will acquire sooner or later, should come from a laboratory where critical observation and accurate measurement are just as much the order of the day now as in the time of its distinguished founder, Alfred Fowler, to whom we remain indebted for so much that is sound in spectroscopic practice.

W. J.

Classical and Modern Physics, a Descriptive Introduction, by HARVEY E. WHITE, Ph.D. Pp. xiii + 712. (London: Chapman and Hall, Ltd., 1940.) 27*s*. net.

So strong and wide is the popular appeal of modern physics that there is now a demand not only for good elementary books on the subject, but also (in the U.S.A. at any rate) for courses of elementary lectures and demonstrations giving the necessary groundwork in classical physics before expounding the modern developments. For some years Professor Harvey White has been giving such a course in the University of California at Berkeley, and has now founded this book upon it. In the preface he explains that his aim has been to give an interesting presentation of the science of physics to those who regard the subject as an essential part of their cultural background, and notes that the tremendous response of Letters and Arts students to the latter part of his course is convincing proof that modern physics can be taught even to the beginner with no previous scientific training. To this appreciation by Professor White's audiences will now be added the thanks of a still wider public for his lucid exposition in this beautiful book.

Each subject treated is started as nearly as possible at the beginning, is developed in such a way that it can be followed by a student with some knowledge of the simplest principles of algebra and plane geometry, and is adequately illustrated with diagrams and photographs. Further, in the presentation of modern physics, the order of its historical development is followed as closely as possible : for each new subject we have, whenever possible, a brief account of the discovery, an experimental demonstration of the phenomenon, a discussion of practical applications and relevant experiments, and finally a brief account of the theory and its experimental confirmation.

After an Introduction consisting of two chapters on optical illusions and units of measurement, there are ten parts, each containing from three to six chapters. Parts I–VI, which occupy rather more than half the volume, deal in a most attractive manner with the older subjects—Mechanics, Properties of Matter, Heat, Sound, Electricity and Magnetism, and Light. Part VII, which is the first to treat of the more modern developments, and is devoted to the phenomena of discharges through gases, consists of three chapters on the discovery of the electron, atoms and the periodic table, and x rays. In the five chapters of Part VIII we have an introduction to the quantum theory and atomic structure, with descriptions of radioactivity, spectra and their classification, and the photo-electric effect. Atomic and nuclear processes are dealt with in the five chapters of Part IX, under the headings of photon collisions and atomic waves, cosmic rays, atomic collisions and nuclear disintegrations, induced radioactivity, and the atomic nucleus. Part X, dealing with astrophysics, consists of three chapters on the sun, the stars and the theory of relativity. Questions and numerical problems are set at the ends of the chapters, and answers are given at the end of the book. Two biographical features are included, which add much to the interest of the book : a selection of footnotes on outstanding men of science and a list of Nobel laureates in physics and the subjects for which the prizes were awarded. A table of stable isotopes is also appended, and a good 16-page index is provided.

The few corrections and suggestions which the reviewer has to offer are being sent to the author with a view to the second edition rather than noted here. The book is heartily recommended to all who wish to derive the maximum enjoyment from their physics. To members of our Physical Society the brief descriptions and beautiful illustrations of the Berkeley cyclotron and the great telescopes in California (in Parts IX and X and the frontispiece) will be of particular interest in the year when the Duddell Medal has gone to Lawrence himself and the 200-inch telescope has formed the subject of the Thomas Young Oration.

W. J.

THE PHYSICAL SOCIETY

MEMBERSHIP

Membership of the Society is open to all who are interested in Physics:

FELLOWSHIP. A candidate for election to Fellowship must as a rule be recommended by three Fellows, to two of whom he is known personally. Fellows may attend all meetings of the Society, are entitled to receive Publications 1, 3, 4 and 5 below, and may obtain the other publications at much reduced rates.

STUDENT MEMBERSHIP. A candidate for election to Student Membership must be between 18 and 26 years of age and must be recommended from personal knowledge by a Fellow. Student Members may attend all meetings of the Society, are entitled to receive Publications 1, 3 and 4, and may obtain the other publications at much reduced rates.

PUBLICATIONS

1. *The Proceedings of the Physical Society*, published six times annually, contains original papers, lectures by specialists, reports of discussions and of demonstrations, and reviews.
2. *Reports on Progress in Physics*, published annually, is a comprehensive review by qualified physicists.
3. *The Catalogue of the Physical Society's Annual Exhibition of Scientific Instruments and Apparatus*. This exhibition is held normally in January of each year.
4. *The Agenda Paper*, issued fortnightly during the session, informs members of the programmes of future meetings and business of the Society generally.
5. *Physics Abstracts* (*Science Abstracts A*), published monthly in association with the Institution of Electrical Engineers, now covers practically the whole field of contemporary physical research.
6. *Electrical Engineering Abstracts* (*Science Abstracts B*), published monthly in association with the Institution of Electrical Engineers, covers in electrical engineering a field similar to that covered by *Physics Abstracts* in pure physics.
7. *Special Publications*, critical monographs and reports on special subjects prepared by experts or committees, are issued from time to time. Some of these special publications and their prices (inclusive of postage) are :

REPORT ON THE TEACHING OF GEOMETRICAL OPTICS (1934). Price: 6*s*. 3*d*.

REPORT ON BAND SPECTRA OF DIATOMIC MOLECULES. By W. JEVONS, D.Sc., Ph.D. (1932). Price: bound in cloth, 21*s*.; in paper covers, 18*s*.

REPORTS OF DISCUSSIONS ON VISION (1932), AUDITION (1931) and PHOTO-ELECTRIC CELLS AND THEIR APPLICATIONS (1930). Prices 6*s*. 4*d*., 3*s*. 8*d*. and 6*s*. 4*d*., respectively.

MEETINGS

At approximately fortnightly intervals throughout each normal session, the Society holds meetings for the reading and discussion of papers, for experimental demonstrations and for special lectures, including the *Guthrie Lecture*, in memory of the founder of the Society, given annually by a physicist of international reputation, the *Thomas Young Oration*, given biennially on an optical subject, and the recently founded *Rutherford Memorial Lecture*. In addition, a Summer Meeting is generally held each year at a provincial centre, and meetings are arranged from time to time jointly with other learned Societies for the discussion of subjects of common interest.

SUBSCRIPTIONS

Fellows pay an Entrance Fee of £1. 1*s*. and an Annual Subscription of £2. 2*s*. Student Members pay only an Annual Subscription of 10*s*. 6*d*. No entrance fee is payable by a Student Member on transfer to Fellowship.

Further information may be obtained from the

ASSISTANT SECRETARY, THE PHYSICAL SOCIETY

1 LOWTHER GARDENS, EXHIBITION ROAD, LONDON, S.W.7